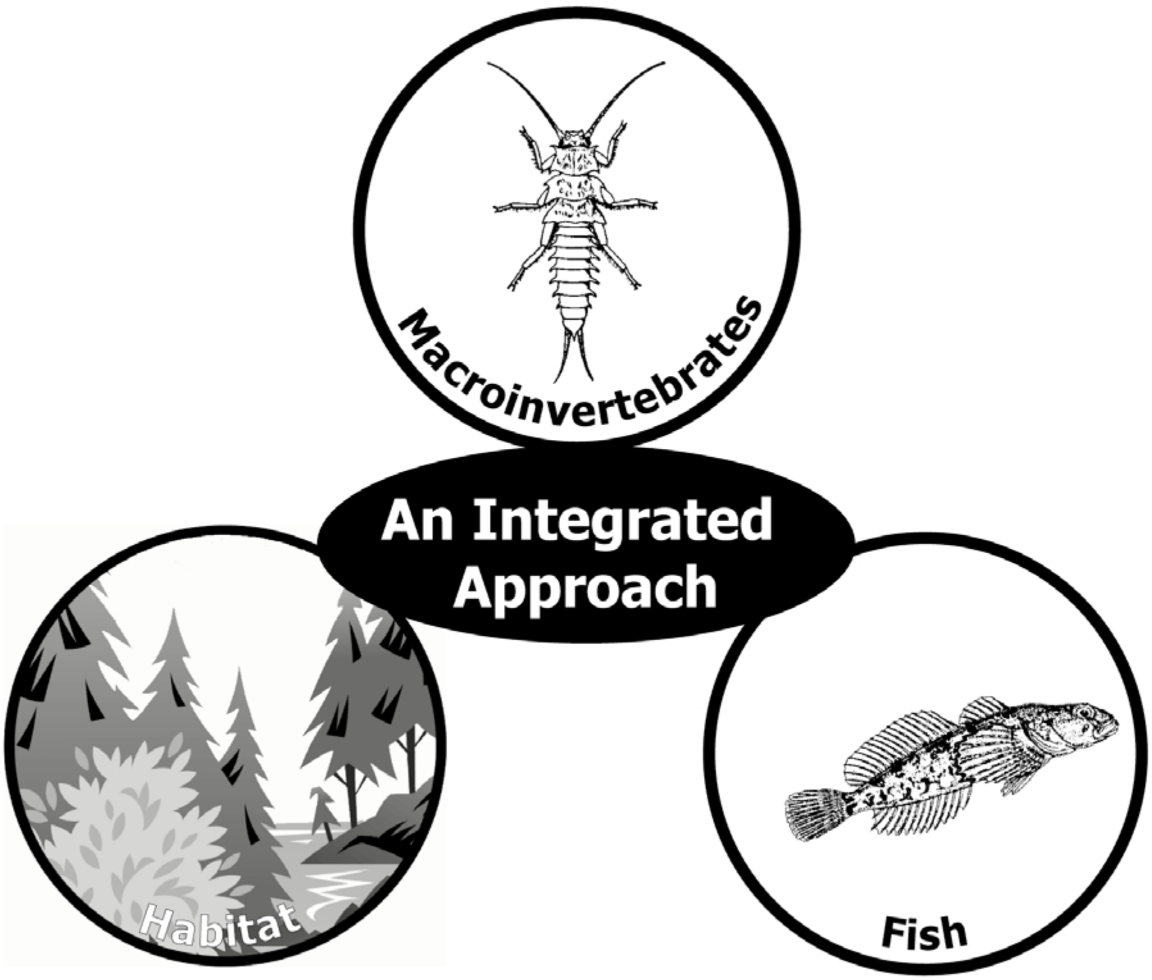
Idaho Small Stream Ecological Assessment Framework



Idaho Department of Environmental Quality

Final – June 2002

Idaho Small Stream Ecological Assessment Framework:

An Integrated Approach

Final
June 2002

IDAHO DEPARTMENT OF ENVIRONMENTAL QUALITY

Acknowledgements

The development of the *Stream Ecological Assessment Framework* could not have been achieved without the dedicated effort of and suggestions from the Stream Assessment Team members. These individuals include: Cindy Barrett, Bill Clark, Lynn Van Every, Cyndi Grafe, Gretchen Hayslip, Chris Mebane, Dave Mosier, Angie Petersen, Steve Robinson, Jack Skille, Bob Steed, Daniel Stewart, and Sean Woodhead.

Special thanks also to Mike Edmondson who managed the Tetra Tech contract for the Stream Macroinvertebrate Index development. He also provided much needed database support by providing data sets for the development of the Stream Macroinvertebrate Index and Stream Habitat Index. Additionally, Mike Edmondson, Henry Navorro, Darren Brandt, and Chris Mebane accomplished considerable data testing and database tool development to evaluate the performance of all the indexes.

We sincerely appreciate the technical peer review provided by Steve Bauer, Christopher Robinson, and EPA. EPA provided DEQ with comprehensive comments from technical and policy perspectives. The EPA reviewers possessed a wide range of expertise including fish biology, ecology, monitoring, program policy and legal. The review included EPA Region 10 Idaho Operations office (Boise, Idaho): Leigh Woodruff. Also, from EPA Region 10 office (Seattle, Washington): Kerianne Gardner, Gretchen Hayslip, Lilian Herger, Curry Jones, Marcia Lagerloef, Theresa Pimentel and Steve Ralph. Lastly, from EPA Headquarters (Washington, D.C.): Susmita Dubey (Office of General Council); Susan Holdsworth, Mike Haire, Chris Faulkner, Christine Ruf (Office of Wetlands, Oceans and Watersheds); Sue Gilbertson, Ed Hanlon, Jennifer Wigal (Office of Science and Technology).

Also, greatly appreciated are the editorial and production support provided by Amy Luft, Pat Jones, and Barbara Mallard.

APPROPRIATE CITATION

Grafe, C.S., ed. 2002. Idaho Small Stream Ecological Assessment Framework: An Integrated Approach. Idaho Department of Environmental Quality; Boise, Idaho.

This entire document can be downloaded from the Idaho Department of Environmental Quality's website at http://www2.state.id.us/deq.

Table of Contents

List of Figures

List of Tables

Acronyms and Symbols

ALUS	Aquatic Life Use Support
ANOVA	Analysis of Variance
BC	Bray-Curtis
BLM	Bureau of Land Management
BM	Blue Mountains
BURP	Beneficial Use Reconnaissance Program
cfs	cubic feet per second
CP	Columbia Plateau
CPUE	Catch per unit effort
C.V.	coefficient of variation
CWA	Clean Water Act
DE	discrimination efficiency
DELT	Deformities, eroded fins, lesions, or tumors
DEQ	Idaho Department of Environmental Quality
DFA	discriminant function analysis
EPA	Environmental Protection Agency
EPT	Ephemeroptera, Plectoptera, Trichoptera
GIS	Geographic Information System
HBI	Hilsenhoff's Biotic Index
HI	habitat index
IBI	Index of Biotic Integrity
IDFG	Idaho Department of Fish and Game
ISU	Idaho State University
m	meter
m^2	square meter
MR	Middle Rockies
NAWQA	National Water Quality Assessment Program
NBR	Northern Basin and Range
NMDS	non-metric multidimensional scaling
NNR	Northern Northern Rockies
NR	Northern Rockies
PCA	principle components analysis
s.d.	standard deviation
SFI	Stream Fish Index
SHI	Stream Habitat Index
SMI	Stream Macroinvertebrate Index
SNR	Southern Northern Rockies
SRP	Snake River Plain
TMDL	Total Maximum Daily Load
USFS	United States Forest Service
USGS	United States Geological Survey
WB	Wyoming Basin
WBAG	Water Body Assessment Guidance
WUM	Wasatch/Uinta Mountains

Abstract

This document describes the Idaho Department of Environmental Quality's (DEQ) ecological assessment approach to determine aquatic life use support in Idaho's small streams. DEQ uses biological indicators, habitat data and numeric water quality criteria to assess aquatic life use support for small streams. The intent of this document is to provide detailed technical information concerning the development and integration of the Stream Macroinvertebrate Index (SMI), Stream Fish Index (SFI), and Stream Habitat Index (SHI) used in the aquatic life use support determination.

DEQ applies the stream ecological assessment approach based on results from three water body size criteria: stream order, width, and depth. In general, the small stream method is applied to water bodies that have an average water body size criteria rating of less than 1.7.

DEQ uses several bioassessment tools or multimetric indexes to limit reliance on just one tool and still ensure direct measurements of aquatic life. DEQ contracted Jessup and Gerritsen with Tetra Tech, Inc. to develop the SMI. Jessup and Gerritsen used sites identified as least impacted and stressed to develop the SMI. The macroinvertebrate data is evaluated within the context of three bioregions: Northern Mountains, Central and Southern Mountains, and Basins. Based on this classification system, Jessup and Gerritsen identified nine significant macroinvertebrate metrics to characterize water quality condition. These SMI metrics include: total taxa, Ephemeroptera taxa, Plecoptera taxa, Trichoptera taxa, percent Plecoptera, Hilsenhoff Biotic Index, percent five dominant taxa, scraper taxa, and clinger taxa.

The Stream Fish Index is a bioassessment tool that directly measures the achievement of the Clean Water Act "fishable" goal. Mebane identified two different sets of metrics to characterize water quality condition for montane-forested and desert basin-rangeland classifications. The rangeland metrics include: percent cold water individuals, Jaccard's community similarity coefficient, percent omnivores and herbivores, percent cyprinids as longnose dace, percent of fish with certain abnormalities (deformities, eroded fins, lesions, and tumors), and catch per unit effort. The metrics in the forested classification comprise: number of cold water native species, percent cold water individuals, percent sensitive native individuals, number of sculpin age classes (unless sample is comprised solely of salmonids), number of salmonid age classes, and catch per unit effort.

DEQ contracted Fore and Bollman with Statistical Design and Rhithron Biological Associates, respectively, to develop the Stream Habitat Index (SHI). They determined that ecoregion groupings provided the most useful classification approach for the SHI. Fore and Bollman tested habitat measures with land use and biological data. They identified ten habitat measures that signaled water quality conditions. The SHI measures include: instream cover, large organic debris, percent fines less than 2mm in wetted width, embeddedness, number of Wolman size classes, channel shape (undercut), percent bank cover, percent canopy cover, disruptive pressures, and zone of influence.

DEQ integrates multiple data types using a rating and averaging approach. Index scores are adjusted to a common scale using a 1, 2, 3 scoring system. The converted scores are then averaged to provide a single score. Average scores greater than or equal to 2 are fully supporting of aquatic life, while scores less than to 2 are not fully supporting.

Chapter 1. Overview

Cynthia S. Grafe[1]

INTRODUCTION

This document describes the Idaho Department of Environmental Quality's (DEQ) ecological assessment approach to determine aquatic life use support (ALUS) in Idaho's small streams. Associated policies and other beneficial use approaches (e.g., recreation, domestic water supply, etc.) are addressed as part of the DEQ *Water Body Assessment Guidance* (Grafe et al. 2002). The intent of this document is to provide detailed technical information concerning the development of the Stream Macroinvertebrate Index (SMI), Stream Fish Index (SFI), and Stream Habitat Index (SHI) used in the ALUS determination for small streams.

The ALUS for small streams is addressed in this document, while the river ALUS is addressed in the *Idaho River Ecological Assessment Framework: An Integrated Approach* (Grafe 2002). It is important to make this distinction, since DEQ uses different monitoring and assessment protocols depending on water body size. Chapter 2 describes the criteria rating and averaging method DEQ uses to differentiate between small streams and rivers.

REGULATORY BACKGROUND

In 1972, Congress passed public law 92-500, Federal Water Pollution Control Act, commonly known as the Clean Water Act (CWA). The goal of this act was to "restore and maintain the chemical, physical, and biological integrity of the Nation's waters" (Water Pollution Control Federation 1987). The act and the programs it generated have changed over the years as experience and perceptions of water quality have changed. It has been amended 15 times, most significantly in 1977, 1981, and 1987. One of the goals of the 1977 amendment was protecting and managing waters to insure "swimmable and fishable" conditions. This goal, along with the 1972 goal to restore and maintain chemical, physical, and biological integrity, relates water quality with more than just chemistry.

The federal government, through the U.S. Environmental Protection Agency (EPA), assumed the dominant role in defining and directing water pollution control programs across the country. DEQ implements the CWA in Idaho while the EPA provides oversight of Idaho's fulfillment of CWA requirements and responsibilities.

[1] Idaho Department of Environmental Quality, 1410 N. Hilton, Boise, ID 83706.

DEQ is charged with providing a consistent water body assessment method using data collected under the Beneficial Use Reconnaissance Project (BURP) and other similar sources. The assessment methods must determine if a water body is supporting or not supporting beneficial uses such as aquatic life. The *Water Quality Standards and Wastewater Treatment Requirements* are the Idaho legally established rules concerning beneficial uses and associated criteria (Rules of the Department of Health and Welfare, IDAPA 58.01.02).[2]

USE OF ECOLOGICAL INDICATORS

The strength of this ALUS assessment framework is the use of ecological indicators. Water quality is evaluated and compared to levels needed for the protection and maintenance of viable communities of aquatic species. Measurements of aquatic assemblages reflect long-term stream conditions more than instantaneous chemical measurements and provide a direct measure of the aquatic life beneficial use. DEQ uses both biological indicators and numeric water quality criteria to assess ALUS. Levels of aquatic life protection and maintenance are evaluated within the context of the Idaho water quality standards goals.

USE OF MULTIMETRIC INDEXES

To evaluate ALUS, DEQ applies multimetric indexes based on rapid bioassessment concepts developed by EPA (Barbour et al. 1999). Measurements of biological or physical habitat conditions known as metrics comprise the indexes. The indexes include several characteristics to gauge overall ecosystem health. The multimetric index value for a sample site is the sum of individual metric scores. Multimetric index scores are unitless, and therefore easily comparable. The index scores from the identified least impacted or reference sites are then used to develop a range of conditions that can be divided into any number of categories indicating different levels of impairment (Barbour et al. 1999). The group of least impacted or reference sites is known as the reference condition and is the benchmark used in the assessment process. DEQ compares the multimetric index scores of sites to this reference condition to determine use support.

The strength of such an approach is the integration of biological, physical, and chemical characteristics of the water body at different scales — individual, population, community, and ecosystem scales (Karr et al. 1986). This integration allows DEQ to detect water quality impairment cost-effectively and furnish this information in an understandable format.

Data used to calculate certain indexes, such as the SFI, may be limited due to sampling resource requirements, endangered or threatened species sampling restrictions, and sampling protocols incompatible with BURP methods. Therefore, DEQ has developed

[2] Henceforth, subsection 3 of regulation within IDAPA 58.01.02 are abbreviated as "WQS.XXX" where XXX is the subsection. For example, "IDAPA 58.01.02.100" is abbreviated as "WQS 100."

several bioassessment tools to limit reliance on just one tool and still ensure direct measurements of aquatic life. The stream ecological assessment framework integrates potentially three multimetric indexes in the ALUS determination: SMI, SFI, and SHI.

Stream Macroinvertebrate Index

DEQ contracted Tetra Tech, Inc. to develop the SMI. Benthic macroinvertebrates are aquatic insects found in the bottom substrate of streams. Jessup and Gerritsen (2002) used sites identified as least impacted and as stressed to develop the SMI. The macroinvertebrate data is evaluated within the context of three classes or bioregions: Northern Mountains, Central and Southern Mountains, and Basins. Jessup and Gerritsen (2002) identified nine significant macroinvertebrate metrics to characterize water quality condition. These SMI metrics include assemblage attributes such as richness, composition, pollution tolerance, diversity, feeding group, and habit. Jessup and Gerritsen also performed a reanalysis of the SMI in the Northern Mountains using a refined data set as well as evaluated ambiguous taxa for all bioregions. Chapter 3 details the data set and methods used to develop the SMI.

Stream Fish Index

The SFI is a bioassessment tool which directly measures the achievement of the Clean Water Act “fishable” goal. Mebane (2002) used sites identified as least impacted and stressed to develop the SFI. He developed two site classes: Montane-Forested and Desert Basin-Rangeland. Mebane (2002) identified two different sets of metrics to characterize water quality condition for forested and rangeland classes. For rangeland sites, six metrics were identified comprising assemblage attributes such as richness, composition, indicator, abundance, and condition. The forest metrics also included richness, composition, indicator, and abundance characteristics as well as reproductive function attributes. Also both classifications incorporate amphibian indicators as a secondary metric. Mebane’s discussion of index development is found in Chapter 4.

Stream Habitat Index

DEQ contracted Statistical Design and Rhithron Biological Associates to develop the SHI. Fore and Bollman (2002) determined that ecoregion groupings provided the most useful classification approach for the SHI. Fore and Bollman (2002) used land use data to evaluate human disturbance gradients in the Snake River Basin and Northern Basin and Range ecoregions. In the Northern/Middle Rockies it was more difficult to develop a disturbance gradient using available land use data, so DEQ professional biologists identified least impacted sites and stressed sites based on observations of human disturbance at the site and in the watershed. Fore and Bollman (2002) also tested habitat measures with fish and macroinvertebrates. Ultimately, they identified ten habitat measures that signaled water quality conditions. Five of these metrics are quantitatively measured, while the other five are field rated using eye estimates. Chapter 5 describes the Fore and Bollman approach to developing the SHI.

DATA INTEGRATION AND REPORTING OF ASSEMBLAGES

To be meaningful to managers and the public, biological data need to be translated into coherent information that conveys the assessment results. The challenge is to interpret and report all the results from different assemblages, particularly when the results are varied or contradictory. DEQ integrates multiple data types by classifying results from the indexes using a 1, 2, 3 scoring system. The converted scores are then averaged to provide a single score that is interpreted for the ALUS determination. Chapter 6 describes the data integration approach and provides an example using actual data to more clearly explain the method.

REFERENCES

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid bioassessment protocols for use in streams and rivers: periphyton, benthic macroinvertebrates, and fish. Second edition. Environmental Protection Agency, Office of Water, Washington, DC.

Clean Water Act (PL 92-500)

Fore, L. and W. Bollman. 2002. Stream habitat index. Chapter 5. *In* Grafe, C.S. (editor). Stream ecological assessment framework: an integrated approach. Idaho Department of Environmental Quality. Boise, Idaho. pp.302.

Grafe, C.S. (editor). 2002. Idaho river ecological assessment framework: an integrated approach. Idaho Department of Environmental Quality. Boise, Idaho. 210 pp.

Grafe, C.S., C.A. Mebane, M.J. McIntyre, D.A. Essig, D.H. Brandt, and D.T. Mosier. 2000. The Idaho Department of Environmental Quality Water Body Assessment Guidance, Second Edition. Idaho Department of Environmental Quality. Boise, Idaho.

Jessup, B. and J. Gerritsen. 2002. Stream macroinvertebrate index. Chapter 3. *In* Grafe, C.S. (editor). Stream ecological assessment framework: an integrated approach. Idaho Department of Environmental Quality. Boise, Idaho. pp.302.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessment of biological integrity in running water: a method and its rationale. Illinois Natural History Survey Special Publication Number 5, Champaign, IL.

Mebane, C.A. 2002. Stream fish index. Chapter 4 *In* Grafe, C.S. (editor). Stream ecological assessment framework: an integrated approach. Idaho Department of Environmental Quality. Boise, Idaho. pp. 302.

Chapter 2.
Water Body Size Criteria

Cynthia S. Grafe[3]

INTRODUCTION

DEQ applies different monitoring and assessment protocols depending on water body size. Since individual perceptions of water body size vary, DEQ developed a consistent method of selecting and applying criteria to distinguish between small streams and rivers.

DEQ examined several water body size criteria and determined that no one criteria could characterize the varying sizes of Idaho streams. Some of the criteria considered were stream order, width, depth, discharge, and drainage area. These criteria were suggested in the literature and by Idaho State University (ISU) (Hughes et al. 1986, Royer and Minshall 1999). This chapter describes the reasons for adopting certain criteria, the application of the criteria, and supporting analysis.

METHODS

Criteria Consideration and Calculation

DEQ considered several criteria to determine water body size. The following is a description of each criterion and the methods used to determine the criteria.

- Stream order – This criterion is often used to determine water body size (Allan 1995) since it is relatively constant. However, with larger water bodies it can be very difficult to calculate the stream order using 1:24,000 topographical maps. For this reason, DEQ followed ISU's protocol (Royer and Minshall 1997) which used the Strahler (1957) method with 1:100,000 Geographic Information System (GIS) hydrography coverage and/or topographical maps. According to ISU, the stream order may be one order less using a 1:100,000 scale (Schomberg, personal communication, 1998). In cases where the water body is extremely large, such as the Snake River, it was assumed that the stream order was seven or greater. ISU only used this *a priori* criterion to distinguish water body size during the development of the RMI.

- Average width at baseflow (m) – This criterion is a measure of water conditions during baseflow when BURP sampling occurs. This is the average wetted width of all measurements taken at the site (n=6). Average width does not discern the

[3] Idaho Department of Environmental Quality, 1410 N. Hilton, Boise, ID 83706.

difference in water body size due to diversions or other water flow regulations. However, ISU did recommend this criterion be used to distinguish water body size.

- Average depth at baseflow (m) – This is an average of all the depth measurements taken at a site (n=approximately 60). These measurements are taken at the transects where macroinvertebrates are sampled. Similar to average width, this criterion assesses conditions during baseflow, but does not necessarily consider water flow regulations. ISU also recommended this criterion be used to distinguish water body size.

- Average greatest depth (m) – This is an average of the three greatest depths in the reach. Originally, DEQ was more concerned about wadability when selecting monitoring protocols for different size water bodies. Specifically, if the water depth was too great to use a Hess sampler per the wadeable stream protocol, then the nonwadeable monitoring protocol would be used. However, DEQ decided that wadability should not be the key criterion for using river protocols. Nonetheless, the average width and depth does take wadability into account and considers if a Hess sampler is appropriate for monitoring.

- Site discharge (cfs) – This is the discharge measured, either by the crew or by a nearby gaging station, on the sampling day. The DEQ protocol is to measure the discharge if the system is fairly wadeable and there is no nearby gage. If DEQ did not take a discharge measurement, then an extrapolation technique was used to determine discharge. There was concern that this criterion would be affected by flow diversions during the sampling period.

- Mean annual site discharge (cfs) – Similar to the site discharge, the mean annual site discharge is determined using data from nearby USGS gaging stations and a similar extrapolation technique. Hughes et al. (1986) suggested using mean annual discharge as a better measure of water body size than stream order. Additionally, DEQ determined that this long-term criterion should generally not be as influenced by flow diversions occurring during baseflow conditions on a particular sampling date.

- Site drainage area (m^2) – This criterion, which measures the drainage area above the site, is calculated using GIS hydrography (1:100,000) and Hydrologic unit codes (HUC) (4th and 5th field) coverages. Site drainage area was also suggested by Hughes (1986) as representative of water body size. DEQ was concerned about using this criterion because flows from similar drainage areas may vary dramatically in southern and northern Idaho due to climate differences.

Criteria Determination

DEQ eliminated some of the considered criteria for several reasons. DEQ decided that discharge should not be used because of influences from flow diversions. Also, this criterion was not recommended by ISU during the development of the River Macroinvertebrate Index (see Grafe 2002). Further, DEQ believed it would be difficult and time consuming to obtain mean annual discharge figures for Idaho water bodies.

DEQ did not select the site drainage area criterion because of the difficulty in calculating this information for all the water bodies and climatic differences between northern and southern Idaho.

DEQ ultimately selected stream order, average width, and average depth. These criteria were recommended by ISU (Royer and Minshall 1999) during the development of the River Macroinvertebrate Index to distinguish among different size water bodies. As noted previously, the formulation of the water body size criteria policy occurred during the development of the river ecological assessment framework. Consequently, DEQ wanted to ensure consistent application of criteria used in the development of river bioassessment tools. Most importantly, DEQ found that the integration of these three criteria seemed to adequately interpret water body size. The addition of the other considered criteria did not seem to significantly change the assignment of water body size classes (Grafe 2002).

Criteria Rating and Assignment of Water Body Size

For bioassessment purposes, DEQ has condensed the ISU size distinctions into two categories: small and large. The criteria and corresponding size categories are located in Table 2-1.

Table 2-1. Water body size categories used to rate each criterion.

Water Body Size Category	Stream Order	Ave. Width at Base Flow (m)	Ave. Depth at Base Flow (m)	Rating
Large	≥5	≥15	≥0.4	3
Small	<5	<15	<0.4	1

DEQ rates each criterion and then averages the rating or score. Through additional analysis, DEQ found that only two size categories, streams and rivers, were necessary to represent small to large water body characteristics for bioassessment purposes. Consequently, DEQ designates water bodies with average scores of greater than or equal to 1.7 as "rivers" while those water bodies scoring less than 1.7 would be classified as "streams" (see Table 2-2).

DEQ chose 1.7 based on the different combinations of rating results. Specifically, if a water body rated twice (1+1) in the small water body size category and only once (3) in the large category, then the total of five would result in an average score rating of 1.67, just below 1.7. Water bodies that have inconsistent scores in the three categories should be further evaluated using additional measures of stream size. The ultimate goal of determining water body size should be to ensure that the proper aquatic life use assessment process (see Section 6) is used. If the water has physical and biological characteristics indicative of a river rather than a stream the assessor needs to use the river assessment process.

Table 2-2. Water body size average score rating categories.

Water Body Class	Average Score Rating
River	≥1.7
Stream	<1.7

EXAMPLE

Tables 2-3 and 2-4 provide an example of the Raft River to illustrate the rating and scoring method. Although the name “Raft River” might conjure up pictures of a large river, this water body is actually a small stream. Table 2-3 shows the stream order, average width, and average depth results. It is important to note that the width and depth measurements were taken during baseflow conditions. Referring to Table 2-1, all the criteria fall within the “small” category and therefore should be rated “1”. Table 2-4 shows the rating for each criterion and the average water body size score, which is “1,” calculated from the data in Table 2-3. Since the average water body size score 1.7 or less, the Raft River is classified as a small stream.

Table 2-3. Example of data used to rate criteria.

Stream	Site I.D.	Stream Order	Average Width (m)	Average Depth (m)
Raft River	1999STWFA041	3	4.3	0.16

Table 2-4. Example of rating each criterion.

Stream	Site I.D.	Stream Order	Average Width	Average Depth	Criteria Average and Score
Raft River	1999STWFA041	1	1	1	1

REFERENCES

Allan, J.D. 1995. Stream ecology: structure and function of running waters. Chapman and Hall, London.

Grafe, C.S. (editor). 2002. Idaho River Ecological Assessment Framework: An Integrated Approach.

Hughes, R.M., Larsen, D.P., and Omernik, J.M. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management. Vol. 10, 5: 629-635.

Royer, T.V. and G.W. Minshall. 1999. Bioassessment methods for Idaho Rivers: validation and summary. Final report to the Idaho Division of Environmental Quality. Department of Biological Sciences, Idaho State University, Pocatello, ID. 30 pp.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. American Geophysics Union Transactions. 38:913-920.

Chapter 3.
Stream Macroinvertebrate Index

Benjamin Jessup[4] and Jeroen Gerritsen[5]

INTRODUCTION

DEQ is developing biological assessment tools for measuring the quality of streams as part of the state's bioassessment program. Through the 303(d) and Total Maximum Daily Load (TMDL) framework outlined in the CWA of 1972 (and revisions of 1977 and 1987 [PL-92-500]), those waters considered to be impaired and threatened must be identified and improved to meet their designated uses. The definition of impairment by natural resource management or regulatory agencies is typically based on attainment or non attainment of numerical water quality standards associated with a water body's designated use (WQS). If those standards are not met (or attained), then the water body is considered to be impaired. Resident biota in a watershed function as continual natural monitors of environmental quality, responding to the effects of episodic as well as cumulative pollution and habitat alteration. Ambient biological surveys are one of the primary approaches to biomonitoring. These surveys, in turn, are used to measure the attainment of biological integrity. The assessment of ecosystem health cannot be done without measuring the attainment of biological integrity goals as directed by the EPA and characterized by the state of Idaho.

The CWA (PL-92-500) has as one of its primary goals the maintenance and restoration of biological integrity, which incorporates biological, physical, and chemical quality. Biological integrity is commonly defined as "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity and functional organization comparable to that of the natural habitat of the region" (Karr and Dudley 1981, Gibson et al. 1996). This concept refers to the natural assemblage of indigenous organisms that would inhabit a particular area if it had not been affected by human activities. This integrity or naturally-occurring structure and function of the aquatic community becomes the primary least impacted condition used to measure and assess water bodies in a particular region.

Careful measurement of the natural aquatic ecosystem and its constituent biological communities can reveal the condition of biological integrity. Several key attributes are measured to indicate the quality of the aquatic resources. Biological surveys establish the attributes or measures used to summarize several community characteristics, such as taxa richness, number of individuals, sensitive or insensitive species, observed pathologies, other biological and ecological elements, and the presence or absence of essential habitat features.

---

[4] Tetra Tech, Inc., 10045 Red Run Blvd., Suite 110, Owings Mills, MD 21117.
[5] Tetra Tech, Inc., 10045 Red Run Blvd., Suite 110, Owings Mills, MD 21117.

Multimetric, invertebrate indexes of biotic integrity, variously called Rapid Bioassessment Protocol (Plafkin et al. 1989, Barbour et al. 1999), Invertebrate Condition Index (Ohio EPA 1989), Benthic IBI (Kerans and Karr 1994) have been developed for many regions of North America and are generally accepted for biological assessment of aquatic resource quality (e.g., Gibson et al. 1996, Southerland and Stribling 1995, Karr 1991). The framework of bioassessment consists of characterizing least impacted conditions from least impacted sites by identifying appropriate biological attributes with which to measure the conditions. These biological conditions within the least impacted sites may be representative of sustainable ecosystem health. The biological condition of any site can be assessed by measuring the specified biological attributes and comparing them to conditions in the least impacted sites.

Biological measurements, called metrics, represent elements of the structure and function of the bottom-dwelling macroinvertebrate assemblage. Metrics change in some predictable way with increased human influence (Barbour et al. 1996). They include specific measures of diversity, composition, and functional feeding group representation and include ecological information on tolerance to pollution. Multimetric indexes, such as the Index of Biotic Integrity, incorporate multiple biological community characteristics and measure the overall response of the community to environmental stressors (Karr et al. 1986, Barbour et al. 1995). Such a measure of the structure and function of the biota (using a regionally-calibrated multimetric index) is an appropriate indicator of ecological quality. Indexes can reflect biological responses to changes in physical habitat quality and water quality, as well as changes at the landscape level (geology, soil, and land uses) to the degree that they affect the sampled habitat.

The objectives of this study were 1) to develop a regionallycalibrated multimetric biological index for Idaho streams using the benthic macroinvertebrate assemblage, 2) to evaluate the index using independent test data and other indexes, and 3) to recommend appropriate applications for the index. Statewide biological stream assessment data from 1994-1996 were used for developing an index. An independent subset of the data were reserved to evaluate (confirm) the index. Results of the analysis were used to make recommendations for improving the state's biological sampling program to achieve more reliable assessments of Idaho streams.

Benthic macroinvertebrate assemblages are the most widely used target organisms in biological assessment, due to several advantages (Barbour et al. 1999).

- Macroinvertebrate assemblages are good indicators of localized conditions because they are relatively sedentary.

- Macroinvertebrates integrate the effects of many short-term environmental variations because most species have a life cycle of several months to one year.

- Degraded conditions can often be detected by an experienced biologist with only a cursory examination of the benthic macroinvertebrate assemblage because macro-invertebrates are relatively easy to identify to family.

- Benthic macroinvertebrate assemblages are made up of species that constitute a broad range of trophic levels and pollution tolerances.

- Sampling is relatively easy, requires few people and inexpensive gear, and has minimal detrimental effects on the resident biota.

- Benthic macroinvertebrates serve as a primary food source for fish.

- Benthic macroinvertebrates are abundant and diverse in most streams.

Metrics fall into categories of taxa richness and diversity, pollution tolerance, feeding groups, habit (mode of locomotion), and reproductive frequency. Each metric can be tested for consistent response to stressors and the most responsive metrics can be cumulatively assessed. Such an index incorporates several metrics, to characterize several response signals at once without giving undue weight to any single metric, producing an average response as an index score. All of the metrics selected for an index have proven responsiveness to impairment and the responses can be explained by ecological mechanisms. This cumulative signal from the biotic community is a reliable and integrated indicator of ecological quality.

In biological assessment, conditions in suspected impaired sites are compared to conditions in least impacted sites. The conditions in least impacted sites are perceived as least impacted conditions, following the concept of the "control" in experimental studies. Increasingly, water management agencies in the United States and abroad are using a regional least impacted condition, consisting of a composite of multiple sites, rather than a single site-specific control (e.g., Barbour et al. 1999, Gibson et al. 1996, Hughes et al. 1986, Reynoldson et al. 1995, Wright 1995, Davis and Simon 1995). Regional least impacted conditions are preferred because (1) they allow extrapolation to sites and areas of similar characteristics to the least impacted sites, (2) regional samples allow more robust estimation of spatial variability among sites and prevent the trivial comparison of paired sites, and (3) a set of regional least impacted sites may be more cost-effective than a paired least impacted site for every assessment (Hughes et al. 1986, Barbour et al. 1999).

To ensure that the signal perceived from the benthos is more sensitive to anthropogenic stressors than it is to natural stressors, the biomonitoring sites are divided into groups of relative ecological homogeneity and comparisons are made within the groups. Across the state, the underlying geology, riparian vegetation, elevation, gradient, stream geomorphology and other natural parameters vary by stream and by region. Biological conditions are expected to vary with some of these natural parameters in the absence of stressors. The framework of ecoregions (Omernik 1987) was developed to delineate areas of relatively similar natural characteristics (Figure 3-1). The assumption that least impacted biological communities are similar within ecoregions was tested. According to the results, ecoregions were split or combined into bioregions - geographic regions with similar biological community structure.

The proposed index of biological condition allows streams in Idaho to be rated according to the similarity of the biological metrics to least impacted streams within bioregions. As such, the conditions to which all samples are compared are the best attainable within a bioregion, not hypothesized "pristine" conditions. The numerical index values can be converted to narrative ratings based on appropriate thresholds. The narrative ratings, ranging from "very good" to "very poor," could become primary indicators of stream conditions for natural resource decision-makers as they consider actions and priorities for the streams. Recommended thresholds, based on index value distributions in the least impacted sites, could be applied as biocriteria in the state's stream assessment program after additional testing and development.

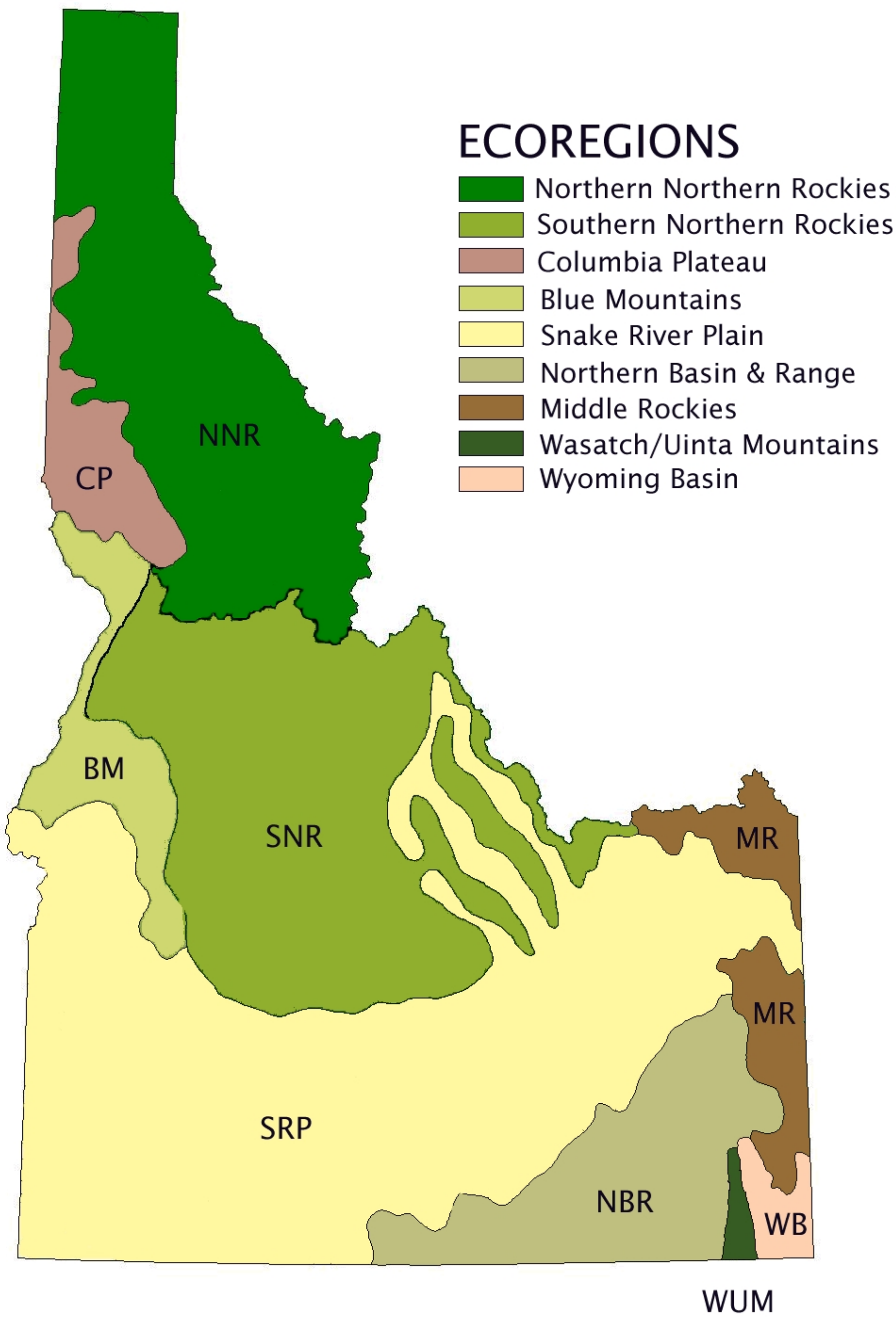


Figure 3-1. Ecoregions of Idaho (Omernik 1987).

METHODS

Biological index development, testing, and application require a stepwise procedure:
1) Definition of site criteria for least impacted and stressed sites,
2) Biological survey of streams,
3) Determination of naturally occurring site classes (bioregional delineations),
4) Testing metric discriminatory ability within bioregions,
5) Scoring candidate metrics and combining metric scores into a regionally calibrated index, and
6) Test index with independent data.

The steps are described below.

Definition of Site Criteria for Least Impacted and Stressed Sites

Professional DEQ biologists developed consensus definitions of least impacted and stressed (most impacted) conditions for Idaho streams (Table 3-1). These non-biotic criteria were intended to identify *a priori* the very best and very worst of sites. Sites must meet all criteria to be designated least impacted, but could be designated stressed by failing several criteria or by severely failing a single criterion (on the judgment of DEQ biologists). Sites not designated as least impacted or stressed were inconsequential to index development procedures.

Table 3-1. Non-biotic variables and criteria used for selecting least impacted sites *a priori*. All of the criteria must be met for a site to be rated "least impacted." "Stressed" sites fail several of the criteria or severely fail any single criterion.

Variable	"Least Impacted" Criteria
Chemical stressors	Likely sources of chemical stress are few (e.g., unbuffered croplands, irrigation returns, active or in-active mining areas, regulated discharges), or if potential sources are present, chemical data shows standards or guidelines are met, and thus effects are unlikely.
Flow modifications	Upstream impoundments absent. Irrigation withdrawal or other diversions absent, or if present, cause minimal disruption to the hydrologic cycle (almost all streams located in the semi-arid basin/lowland ecoregions will have some water withdrawals).
Sedimentation	Causes for anthropogenic sediment increases not apparent (e.g., crop or road gullies, livestock bank trampling, mass wasting). No field notes of highly turbid conditions. Channel substrate of less than 50% fine sediments (measured as bankfull). No "poor" qualitative cobble embeddedness estimates (75%). No "poor" habitat ratings for bank stability.

Variable	"Least Impacted" Criteria
Habitat structure complexity	Substrate is heterogeneous and site contains greater than 30% stable instream fish cover.
Channel complexity	Mixture of pool, glide, riffle, and run habitat types. Longitudinal habitat distribution rated sub-optimal or optimal. No "poor" habitat ratings for channel shape.
Shoreline/channel modifications	Evidence of artificial bank armoring, channel straightening, vegetation removal, or other disturbances absent or minimal.
Riparian vegetation	Riparian growth is extensive and old. It occurs all along the shoreline and is capable of shading the stream and buffering human influences. It overhangs the stream or deposits large woody debris. Bank vegetation protection >70% (sub-optimal or above) or canopy cover shading >25%. No "poor" habitat ratings for bank vegetative protection.
Riparian structure	The riparian vegetative structure has a canopy, understory, and ground cover (trees, shrubs, and ground cover).
Land use	Roads, logging, construction, farming, grazing, and other land uses that disturb the natural vegetation and soils are absent, infrequent, or do not impinge on the riparian zone. Evidence of unnatural bank failures, trampling, excess runoff, or irrigation returns absent or minimal. The habitat variables "disruptive pressures" and riparian "zone of influence" should be rated optimal (9-10) (see DEQ BURP field form).

Modified from Hughes (1995)

Biological Survey of Idaho Streams

In 1994, DEQ finalized its methodology for stream reconnaissance (DEQ 1996). Fifteen core parameters are measured at each site, including physical characteristics, habitat (Hayslip 1993), benthic macroinvertebrates (Clark and Maret 1993), and fish (Chandler et al. 1993). Benthic macroinvertebrates are collected with a Hess sampler (500 micrometers [µm] mesh) in three riffle habitat units. The three samples are composited in the laboratory, and the first 500 individuals are identified to the lowest practical level or to a prescribed level (e.g., Chironomidae were identified at the family level). Subsamples meet or exceed the 500 individual target depending on subsampling rules regarding minimum and maximum effort. A subsample with fewer than 500 individuals constitutes the entire sample. Since 1994, several thousand stream sites have been sampled using these protocols.

Data used for index development included benthic macroinvertebrate relative species abundances from riffle samples, location information, physical measurements, and habitat assessments over three summer sampling periods, from 1994 to 1996. The selection of streams for the biomonitoring program was not random, but targeted towards streams of interest to the state because of exceptionally high quality or suspected anthropogenic impacts, among other reasons (DEQ 1996). Data were managed in a relational database for efficient storage and querying.

For certain steps of index development, sites were divided randomly into calibration and test sets. Site classification and final index evaluation used all the data (calibration and test). Approximately 65 percent of the samples were assigned to the calibration set and were used to determine metric discrimination and select metrics for the multimetric index. The remaining 35 percent of the data were used to test the index performance. Proportional representation of each ecoregion in Idaho was maintained in the calibration and test data sets as far as possible.

Determinating Naturally-occurring Site Classes (Bioregional Delineations)

Detecting changes in the biological assemblage due to human effects must take into account inherent differences due to natural factors. Natural variability in the macroinvertebrate assemblage may result from natural variability in the physical and chemical site characteristics across a geographic range. Much of the natural variability can be accounted for by dividing the area into ecological regions. Ecoregions (Omernik 1987) are delineations of areas with similar climate, geology, soils, vegetation, topography, and hydrology. Ecoregions have been accepted as a geographic framework for delineating regions of relatively homogeneous natural conditions (e.g., Barbour et al. 1996).

Two primary classification techniques, ordination and comparison of metric distributions, were used to justify separating or combining data from ecoregions into larger regions of relative biological homogeneity (bioregions). All of the least impacted sites were used in site classification. Stressed sites were not used because they are not truly representative of natural biological conditions. Ecoregions with few samples or that showed inconclusive groupings using the two techniques were combined with ecoregions of similar natural characteristics as determined by state biologists.

Alternative classification schemes were examined with multivariate ordination of the least impacted sites based on their species composition, following methods outlined in Jongman et al. (1987) and Ludwig and Reynolds (1988). Ordination is a category of methods for reducing the dimensionality of multivariate information (many species in many sites) by placing sites or species in an order. The first ordination method we used is non-metric multidimensional scaling (NMDS) using the Bray-Curtis (BC) dissimilarity coefficient. The BC coefficient contrasts relative taxa abundances between samples according to the formula:

$$BC = 1 - \left({2W}/{(A+B)} \right) ,$$

where W is the sum of common taxa abundances and A and B are the sums of taxa abundances in individual sample units. A pair of samples with identical taxa abundances would have a BC coefficient of 0 and a pair of samples with no taxa in common would have a BC coefficient of 1. This ordination method has been shown to be robust for ordination of species composition (e.g., Kenkel and Orloci 1986, Ludwig and Reynolds 1988), and has been used successfully for classification of stream communities (e.g., Barbour et al.1996; Reynoldson et al.1995).

The site-by-site matrix of BC dissimilarity coefficients was used in the NMDS ordinations (McCune and Mefford 1997, Kruskal 1964). An acceptable ordination should have a stress coefficient (measuring the goodness-of-fit of the ordination to the original data) of less than 20 percent. Stress is lowered as additional dimensions are allowed in the ordination, and three axes are commonly required. The final NMDS configuration was plotted (as a scatterplot in two dimensions) to identify groupings of sites with similar taxa composition (low BC dissimilarity). When plotted points are labeled by site characteristics (e.g., ecoregions) the association between taxa composition and site characteristics can be visualized. Ecoregion groupings that overlap in the ordination plots could be combined into bioregions for subsequent analysis.

A second ordination using metric values in a principle components analysis (PCA) showed groupings of sites with similar metric values. Prior to analysis, all metrics were examined for normality using normal probability plots. Transformations were applied as needed (log or arcsin square-root functions) to normalize the metric distributions. All metrics are entered into the analysis as redundancy does not affect the ordination. NMDS and PCA ordinations proceeded using PC-ORD software (McCune and Mefford 1997).

The second technique used to discern bioregional delineations was a comparison of box and whisker diagrams of metric distributions from least impacted sites. Similar distributions of metrics (medians, inter-quartile ranges, and overall ranges) between ecoregions indicate similar biotic assemblages and justify aggregation of ecoregions into a single bioregion. Likewise, differences in distributions suggest distinct bioregions.

Testing Metric Discriminatory Ability within Bioregions

Metrics were included in the analysis based on successful performance in previous studies (Stribling et al. 1998, Gibson et al. 1996, Barbour et al. 1996), including an unpublished study in Wyoming (Stribling et al. 2000). The metrics fall into seven categories: taxonomic richness, composition, pollution tolerance, diversity, feeding group, habit, and voltinism. Metrics calculated from the Idaho data are given in Table 3-2. The general ecological meanings associated with each category are discussed below.

Table 3-2. Definitions of candidate macroinvertebrate metrics and predicted direction of metric response to increasing perturbation (modified after Barbour et al. 1999).

Metric	Definition	Predicted Response to Increasing Perturbation
Richness measures		
Total taxa	Number of distinct taxa in the macroinvertebrate assemblage	Decrease
EPT[1] taxa	Number of taxa in the insect orders Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies)	Decrease
Ephemeroptera Taxa	Number of mayfly taxa	Decrease
Plecoptera Taxa	Number of stonefly taxa	Decrease
Trichoptera Taxa	Number of caddisfly taxa	Decrease
Diptera taxa	Number of “true” fly taxa	Variable
Composition measures		
% EPT	Percent of the sample that is mayfly, stonefly, and caddisfly larvae	Decrease
% Ephemeroptera	Percent of sample that is mayfly nymphs	Decrease
% Plecoptera	Percent of sample that is stonefly nymphs	Decrease
% Trichoptera	Percent of sample that is caddisfly larvae	Decrease
% Elmidae	Percent of sample that is elmid beetle larvae or adults	Decrease
% Hydropsychidae	Percent of sample that is net spinning caddisfly larvae	Variable
% Diptera	Percent of sample that is “true” fly larvae	Increase
% Diptera (non-chironomid)	Percent of sample that is “true” fly larvae, but not of the midge family	Increase
% Chironomidae	Percent of sample of the midge family of flies	Increase
% non-insects	Percent of sample that is not insects	Increase
Pollution Tolerance Measures		
Intolerant taxa (0-1)	Taxa richness of those organisms considered to be most sensitive to perturbation (with a tolerance value of 0 or 1)	Decrease
Intolerant taxa (0-3)	Taxa richness of those organisms considered to be sensitive to perturbation (with a tolerance value of 0 to 3)	Decrease

Metric	Definition	Predicted Response to Increasing Perturbation
% Tolerant	Percent of sample considered to be tolerant of various types of perturbation (with tolerance values of 7 to 10)	Increase
Hilsenhoff Biotic Index	Abundance-weighted average tolerance of organisms to pollution. Originally designed to evaluate organic pollution (Hilsenhoff 1987).	Increase
Diversity Measures		
Shannon-Wiener Index	A measure of the heterogeneity of the community or of the diversity of dominant taxa (Shannon and Weaver 1949)	Decrease
Simpson's Index	The probability of randomly and independently selecting the same taxa from the sample twice (Simpson 1949)	Increase
% Dominant taxon	Percent of sample in the single most abundant taxon. Also calculated as dominant two, three, five, or 10 taxa.	Increase
Feeding Measures		
Scraper taxa	Number of taxa that scrape periphyton from substrates	Decrease
% Scrapers	Percent of sample that are scrapers	Decrease
% Collectors	Percent of individuals that scavenge organic matter	Variable
% Predators	Percent of the sample that are predators but not omnivores	Variable
Habit Measures		
Clinger taxa	Number of taxa that have fixed retreats or adaptations for attaching surfaces in flowing water	Decrease
% Cclingers	Percent of sample that are clingers	Decrease
% Cclingers of insects	Percent of insects in sample that are clingers	Decrease
Voltinism Measures		
Semi-voltine taxa	Number of taxa that have aquatic life cycles lasting more than one year	Decrease
% Semi-voltine	Percent of sample that are semi-voltine	Decrease

[1]EPT: Ephemeroptera, Plectoptera, Trichoptera

Taxa richness metrics. High taxa richness usually correlates with increased health of the assemblage and suggests that niche space, habitat, and food sources are adequate to support the survival and propagation of many species. "Total taxa" measures the overall variety of the macroinvertebrate assemblage. No identities of major taxonomic groups are derived from the total taxa metric, but the elimination of taxa from a naturally diverse system can be readily detected. Subsets of "total" taxa richness are also used to accentuate key indicator groupings of organisms, such as mayflies, stoneflies, and caddisflies.

Composition measures can be characterized by two classes of information: identity and relative abundance. Identity is the knowledge of individual taxa and associated ecological patterns and environmental requirements (Barbour et al. 1995). Key taxa (i.e., those that are of special interest or ecological importance) and their abundance in the targeted assemblage provide information regarding suitability of stream conditions for growth and reproduction.

Pollution tolerance measures characterize the relative sensitivity of the assemblage to perturbation. They measure numbers or percent composition of pollution tolerant and intolerant taxa (Barbour et al. 1995). Tolerance is generally non-specific to the type of stressor. The Hilsenhoff Biotic Index (HBI) (Hilsenhoff 1987, 1988) was originally oriented toward detection of organic pollution using insects. Taxa tolerance values have since been adjusted to account for regional variation in taxa distributions and stressor types and to include the non-insect taxa. The HBI is an abundance-weighted average tolerance:

$$HBI = \sum \frac{n_i * tv_i}{N} \quad ,$$

where n_i is the number of individuals in the i^{th} taxa, tv_i is the tolerance value assigned to the i^{th} taxa, and N is the number of individuals in the sample with known tolerance values. The richness of intolerant taxa can be calculated at different tolerance levels: taxa with tolerance values less than 4, taxa with tolerance values less than 2, etc.

Diversity metrics are good indications of the ability of the ecosystem to support varied taxa evenly. Several diversity indexes, which are measures of information content and incorporate both richness and evenness in their formulas, may function as viable metrics in some cases, but are usually redundant with taxa richness and percent dominance (Barbour et al. 1996). The most common of these is the Shannon-Wiener index (Shannon and Weaver 1949):

$$H' = -\sum_{i=1}^{s} \left(\frac{n_i}{N} * \log_2 \left(\frac{n_i}{N} \right) \right) \quad ,$$

where n_i is the number of individuals in the i^{th} taxa, and N is the total number of individuals in the sample. The base of the logarithm is most commonly 2, to reflect the binary basis of species presence/absence, but any other base could also be used. Percent dominance metrics evaluate the degree to which conditions favor a single taxon or few taxa.

Feeding measures encompass functional feeding groups and provide information on the balance of feeding strategies (food acquisition and morphology) in the benthic assemblage. Feeding groups include scrapers, collectors, and predators. Specialized feeders, such as scrapers, are more sensitive organisms and are thought to be well represented in healthy streams. Generalists, such as collectors, have a broader range of acceptable food materials than specialists (Cummins and Klug 1979), and thus are more tolerant to pollution that might alter availability of certain food.

Habit (mode of locomotion) measures are those that denote the mode of existence and consist of morphological adaptations for maintaining position and moving about in the aquatic environment (Merritt and Cummins 1996). Habit categories include movement and positioning mechanisms such as skaters, divers, swimmers, clingers, sprawlers, climbers, and burrowers. The clingers are adapted to life in running waters and are sensitive to hydrologic perturbation, habitat disturbance, and other pollutants.

Voltinism measures. The percentage and richness of animals having aquatic life cycles longer than one year were calculated. A community where such semi-voltine organisms are well represented indicates that environmental conditions are relatively stabile, with perturbations that are either infrequent or mild relative to the organism's parameters for survival. A community in which no organisms require long residence times for maturation indicates that perturbations disrupt maturation or reproduction.

Using calibration data only, individual metrics were tested for separation of values between least impacted and stressed sites. The metrics exhibiting differing distributions were considered for inclusion in a multimetric index. The discrimination efficiency (DE) was used as the performance measure in this evaluation. It was calculated as the percentage of metric values in stressed sites that were worse than the worst quartile (25^{th} or 75^{th} percentile) of the least impacted metric values. The judgment of better and worse metric values required an understanding of the ecological mechanisms by which stressors influence biological metric values. Metrics were evaluated within bioregions. Consistency of response in all bioregions (relative degree of separation and trend of response) was a prerequisite for a metric's inclusion in the index.

Scoring Metrics and Combining Metric Scores into a Regionally-Calibrated Index

The purpose of an index is to provide a means of integrating information from the various measures of biological attributes (metrics). Metrics vary in their scale; they are integers, percentages, or dimensionless numbers. Prior to developing an integrated index for assessing biological condition, it is necessary to standardize candidate metrics via a transformation to unitless scores. The standardization assumes that each metric has the same importance in the index (i.e., they are weighted the same).

A continuous scoring strategy was used (Barbour et al. 1999), which rates the metric values on a percentage scale from the worst possible value to the optimal value. In this way, all metrics can score between 0 (worst) and 100 (optimal). To minimize the influence of potentially non-representative outliers, the 95^{th} percentile of the data was considered optimal. Metrics that increase with increasing perturbation (such as percent dominant taxon) score best at the 5^{th} percentile and worst at the maximum value recorded in the entire data set. Some of the metrics had skewed distributions; they contained many low values with few higher values (e. g., percent Elmidae, percent non-insects). Decreasing skewness equalizes a metric's contribution to the index compared to unskewed metrics. The transformation used was the arc-sine square root function, which has greatest effect on percentage values less than 15 percent and greater than 85 percent.

The index was calculated as an average of the included metric scores. Alternative indexes were formulated from the metrics with strong and consistent discriminatory ability in all bioregions. The goal of alternative index formulation was to identify the index that: 1) included responsive metrics from all of the metric categories, 2) included metrics that were not redundant, and 3) gave the greatest separation between least impacted and stressed index scores. By including metrics from all categories, the index will incorporate diverse ecological information and may be responsive to a broad range of stressors. Redundant metrics in an index will bias the index towards the common response mechanisms (Barbour et al. 1992). To avoid redundant information in the index, correlation analysis (Pearson product-moment) was performed on metrics from all samples. Any metrics with a correlation coefficient greater than 0.9 or less than -0.9 were considered redundant and were not used together in any index formulations. Metrics with correlation coefficients greater than 0.8 were used together only when necessary to represent information from all metric categories. Each alternative index was tested for its ability to discriminate between least impacted and stressed sites using the index discrimination efficiency within each bioregion.

Because DEQ is interested in comparing the proposed macroinvertebrate index with indexes previously proposed or applied within the state, additional index formulations were tested. Metrics used in the River Macroinvertebrate Index (Royer et al. 2001), Macroinvertebrate Biotic index (described in Mebane 2000), Benthic Index of Biotic Integrity (Kerans and Karr 1994), and the analysis of Robinson and Minshall (1998) were scored and combined for side-by-side analysis with the SMI.

Testing the SMI with Independent Data

The data put aside for testing the SMI was not used in its development. Rather, the SMI was applied to this test data and the discriminatory ability was again evaluated. The same SMI thresholds established with calibration data (the 25th percentile of least impacted SMI scores) were used to calculate the percentage of “correct” site classifications of least impacted and stressed test sites. Some reduction in discriminatory ability may be expected using test data instead of the calibration data. Drastic reduction in discriminatory ability of the index in the test data would warrant reexamination of the index components.

In final steps of SMI development, the calibration and test data were re-combined and the SMI was recalibrated using the entire data set. Scoring formulas and index statistics are presented to aid natural resource managers in application of the SMI. SMI threshold values are suggested to allow narrative rating of sites into categories of biological integrity.

RESULTS AND DISCUSSION

Database Characteristics

DEQ personnel collected 1,758 valid samples from 1,440 biomonitoring sites during the summer index periods of 1994, 1995, and 1996. Using non-biological stream condition criteria (see Table 3-1), 150 least impacted and 145 stressed samples were identified. Samples that included benthic macroinvertebrate and habitat data were considered valid. The sites were located throughout the eight ecoregions of Idaho though the numbers or densities of sites per ecoregion were not equivalent. For development and confirmation of the multimetric index, the 295 samples were randomly divided into calibration and test sets. Sample sizes for the least impacted and stressed calibration and test data are presented in Table 3-3.

Table 3-3. Sample sizes in the ecoregions of Idaho by *a priori* stream condition (least impacted or stressed) and data set assignment (calibration or test).
The Northern Rockies ecoregion was divided into northern and southern portions.

Ecoregion	Least Impacted Calibration/Test	Stressed Calibration/Test
Northern Northern Rockies	17 / 9	13 / 7
Southern Northern Rockies	29 / 15	10 / 7
Middle Rockies	11 / 6	7 / 3
Blue Mountains	11 / 4	5 / 2
Wasatch and Uinta Mountains	3 / 1	0 / 0
Snake River Basin	16 / 8	35 / 21
Northern Basin and Range	9 / 5	16 / 8
Columbia Plateau	2 / 1	4 / 2
Wyoming Basin	2 / 1	3 / 2
Totals	100 / 50	93 / 52

Site Classification

Idaho streams lie within eight ecoregions (Figure 3-1): the Wyoming Basin, the Snake River Basin (SRB), the Northern Basin and Range (NBR), the Columbia Plateau (CP), the Middle Rockies (MR), the Northern Rockies (NR), the Blue Mountains (BM), and the Wasatch-Uinta Mountains (WUM). Preliminary indications that the Northern Rockies could have distinct biological characteristics between the northern and southern portions suggested division of the ecoregion at the watershed between the Clearwater River drainage to the north and the Salmon River drainage to the south. The WB, CP, and WUM had less than five least impacted sites. Because data from such small samples give unreliable signals, these ecoregions were grouped with similar ecoregions using the best professional judgment of the authors and DEQ staff.

The emphasis on ecoregions as the primary grouping variable comes from the understanding that the ecoregions incorporate many forms of ecological information. Climate, geology, soils, topography, vegetation, and hydrology are all considered in the ecoregion framework

(Omernik 1987). Examination of each physical habitat variable separately does not usually result in more clearly defined bioregional groupings. This assumption was investigated during ordination analysis.

Distributions of metric values in least impacted sites were plotted by ecoregion to detect similarities that would suggest a bioregional scheme (Figure 3-2). Few metrics had distributions that showed distinct differences between ecoregions. The northern Northern Rockies appeared unique in the average and maximum numbers of all taxa, Plecoptera taxa, and intolerant taxa. These and other metrics in the northern Northern Rockies were highly variable. Metric distributions among the other mountainous ecoregions were similar to each other, but they were different relative to non-mountainous ecoregions. This was especially evident in the pollution tolerance metrics HBI and intolerant taxa, which indicate that intolerant taxa are more diverse and abundant in the mountainous regions. The WUM distributions aligned with non-mountainous ecoregions, but this distribution represents only four samples and bioregional groupings based on this small sample would be insubstantial. The WUM was grouped with other mountainous regions based on similar physical habitat conditions (Omernik and Gallant 1986). Likewise, the WB metric distributions resembled some mountainous ecoregion distributions, but the small sample size (n = 3) and comparative physical habitat conditions warrant grouping the WB with other non-mountainous ecoregions. Thus, the bioregional scheme we suggest after reviewing these metric distributions includes three bioregions; 1) the Northern Mountains, 2) the Central and Southern Mountains, and 3) the Basins (Table 3-4, Figure 3-3).

NMDS ordination of relative taxa abundance (Figure 3-4) and PCA ordination of metrics (Figure 3-5) revealed that the Northern Mountains and the Basins are somewhat distinct from each other, whereas the other Southern and Central Mountain samples are dispersed throughout the ordination space. The NMDS ordination illustrates the similarity of relative abundances of taxa in the samples (Figure 3-4). The taxa groups (mostly families) that are most common show the strongest correlations to the ordination axes. These taxa include oligochaetes, chironomids, simuliids, baetids, heptageniids, rhyacophilids, and elmids. Stream gradient is correlated with the same axis as oligochaetes, heptageniids, and rhyacophilids. Latitude is correlated with the same axis as chironomids, with baetids at the other end of the axis. The Northern Mountains separate from the Basins on a diagonal axis, with greatest differences noted in the Plecoptera family abundances.

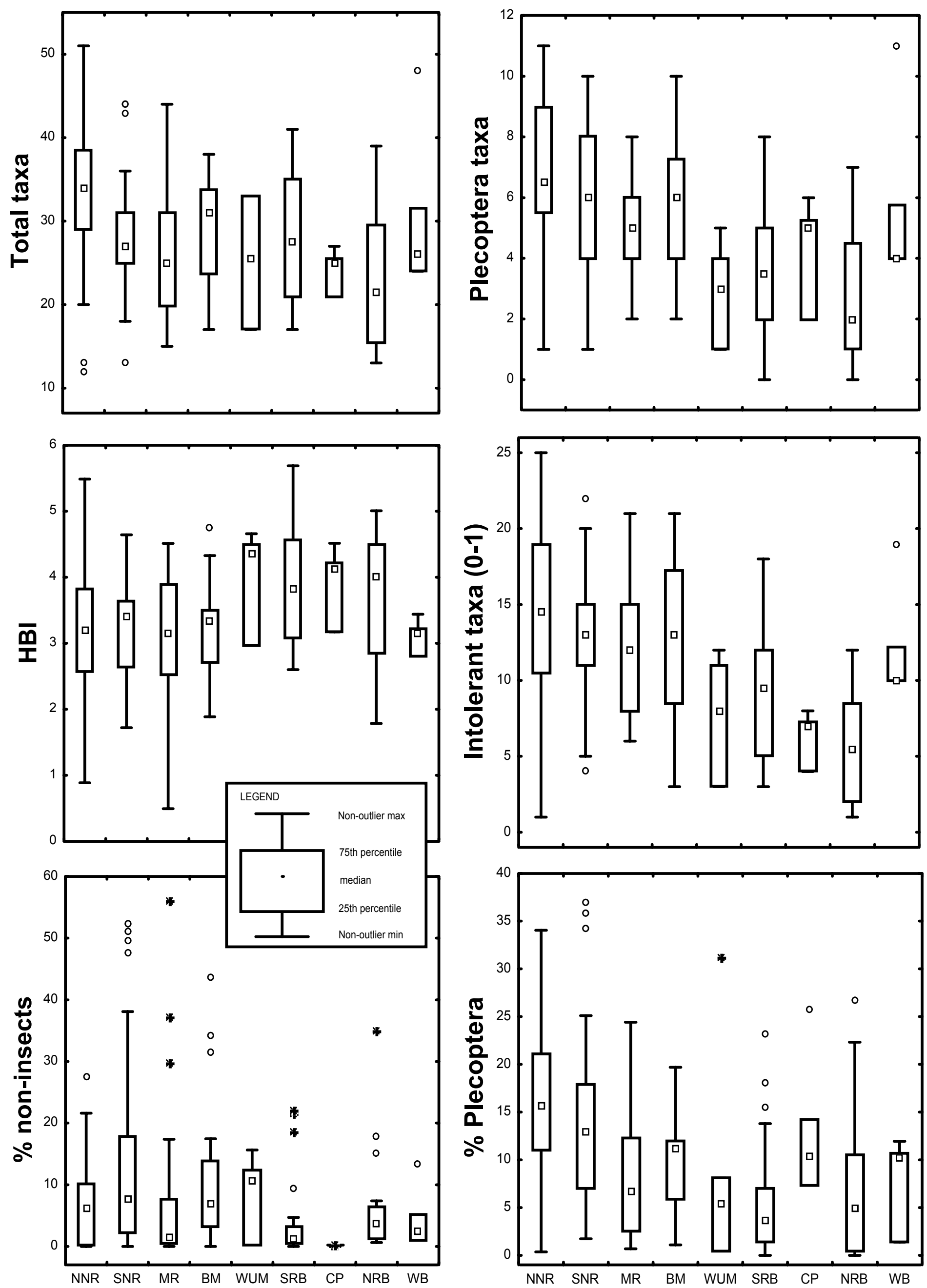


Figure 3-2. Distributions of selected metrics for sites, by least impacted ecoregions of Idaho.
The two divisions of the Northern Rockies are abbreviated NNR and SNR for the northern and southern portions. The WUM, CP, and WB ecoregions have fewer than five least impacted samples.

Table 3-4. Final site classification; three bioregions and the ecoregions they include.

Northern Mountains	Central and Southern Mountains	Basins
Northern Northern Rockies	Southern Northern Rockies	Snake River Basin
	Blue Mountains	Northern Basin and Range
	Middle Rockies	Columbia Plateau
	Wasatch and Uinta Mountains	Wyoming Basin

In the PCA ordination, the first two axes explained 47 percent of the variance (Figure 3-5). The first axis shows strong association with richness related metrics and the second axis is strongly associated with composition (percentage) metrics (Table 3-5). Samples with greater taxa richness in most taxa groups are on the left of the plot, where many of the Northern Mountain samples appear. The second axis generally separates samples with higher % EPT and intolerant individuals (bottom) from samples with higher % Diptera (top). The third axis explained another 10% of the variance and was associated with % Elmidae and % semi-voltine taxa.

These ordinations do not show conclusive evidence for bioregional groupings based on ecoregions. In general, the Northern Mountains have more taxa and more intolerant taxa than the Basins, though many of the taxa abundances are similar. The Central and Southern Mountains cover the ordination space, implying high variability in taxa abundances and metric values. Bioregions can only be defined in conjunction with metric distribution comparison and the expertise of local biologists. Sample sizes in the bioregions are displayed in Table 3-6.

Prevailing physical characteristics and habitat assessment scores in the ecoregions support grouping mountainous and non-mountainous regions, and separating out the northernmost Rockies. Land surface forms, potential natural vegetation, and land uses in the mountains are distinct from those in the plains (Omernick and Gallant 1986). The mountainous terrain supports coniferous forests with timber production and forest grazing as the primary land uses. The basins by contrast are mostly flatter with some tablelands, hills, and low mountains. They support prairie grasses and sagebrush that is grazed or irrigated and cultivated. The basins are generally drier than the mountains as well, receiving less than 25 inches of precipitation per year.

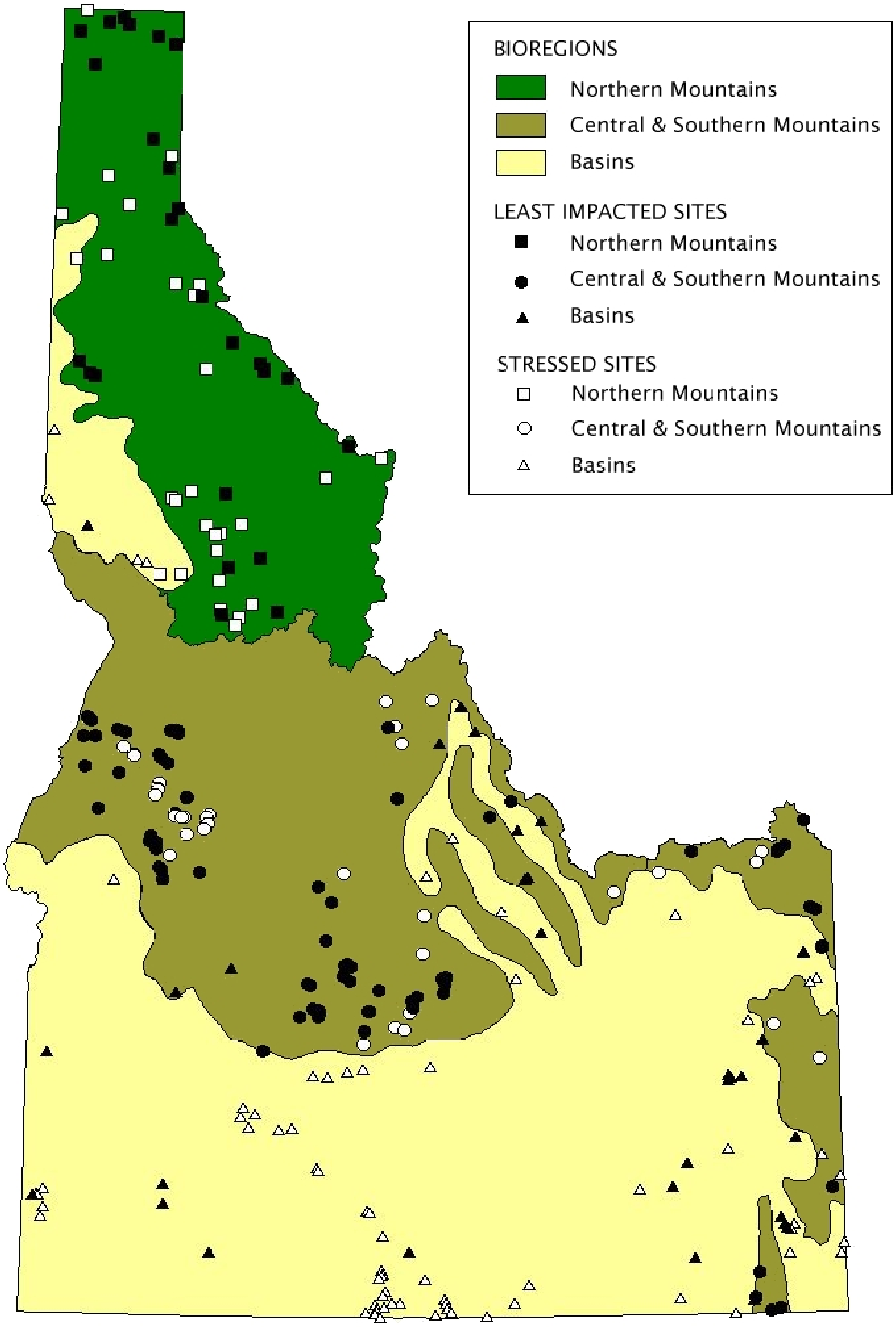


Figure 3-3. Geographic distribution of least impacted and stressed biomonitoring sites in the bioregions of Idaho.

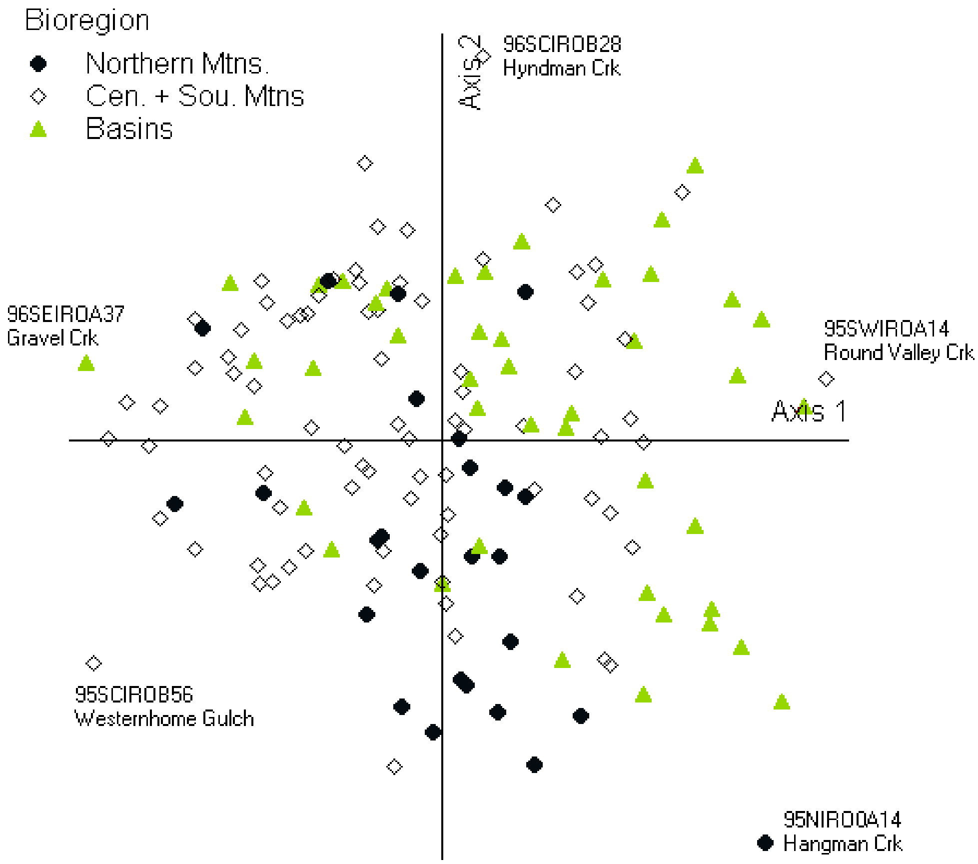


Figure 3-4. NMDS ordination of family level relative abundance for 150 least impacted samples.

Three bioregions are designated and outliers are labeled. In general, samples with greater percentages of oligochaetes, heptageniids, and rhyacophilids are on the left of the plot; chironomids are in the lower right; and baetids and elmids are in the upper right. Oligochaete-dominated samples separate from the others on the third axis (not shown), as do the elmids. Stream gradients are somewhat steeper in samples at the left of the plot.

The northern Northern Rockies and Blue Mountains are lower in elevation than the southern Northern Rockies, Middle Rockies, or Wasatch/Uinta Mountains (Figure 3-6). Compared to the non-mountainous regions, stream gradients are steeper and substrate particle sizes are larger in the mountains. The Middle Rockies ecoregion has more fine particles in its substrate composition than the other mountainous regions. The Snake River Basin has many large streams (by discharge), and the Columbia Plateau is dominated by small streams.

Habitat features in the least impacted sites of the ecoregions were assessed as scores (Table 3-7). The “channel shape” median scores were noticeably higher in the Central and Southern Mountains bioregions than either the Northern Mountains or the Basins bioregions. Other median scores were similar among ecoregions, though the Columbia Plateau and Northern Basin and Range scores were often lower. The highest total habitat scores (total of all habitat feature scores) were in the southern Northern Rockies, Wasatch/Uinta Mountains, and Blue Mountains.

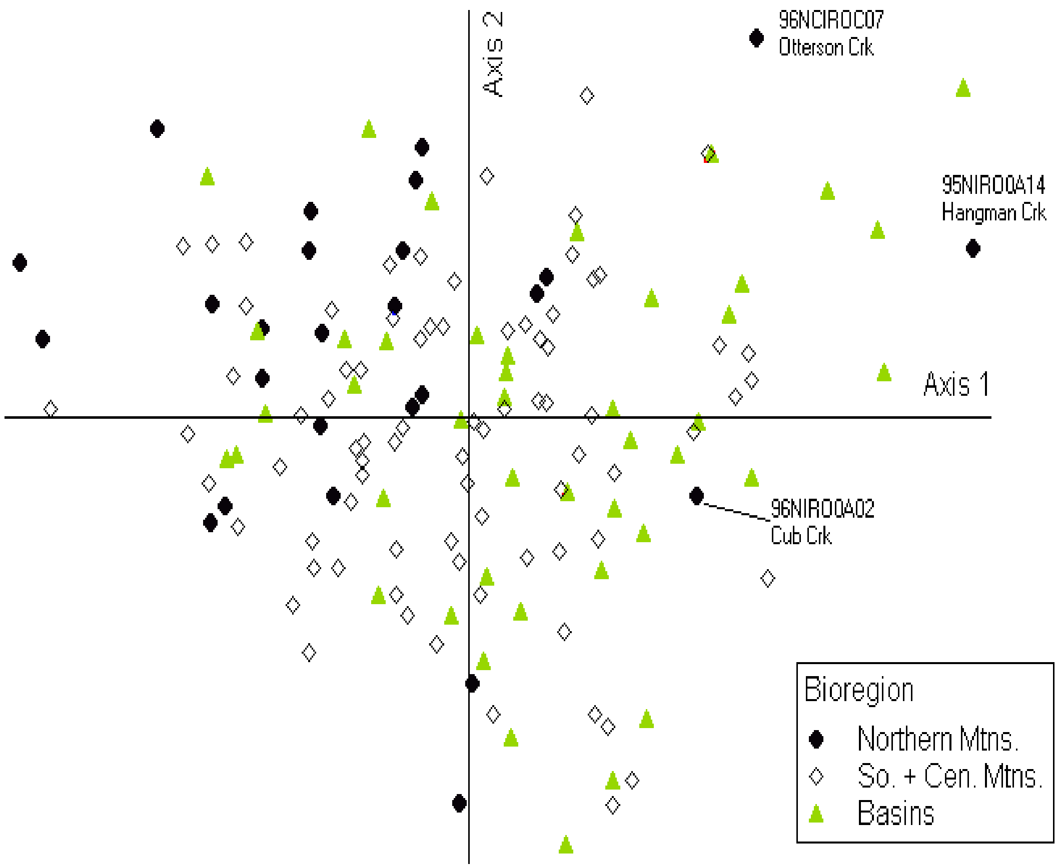


Figure 3-5. PCA ordination of 150 reference sites designated by bioregion, showing the first two axes. Thirty-six metrics with normal or normalized distributions went into the PCA. Samples with greater taxa richness in most taxa groups are on the left of the plot. The second axis generally separates samples with higher percent EPT and intolerant individuals (bottom) from samples with higher percent Diptera (top). Outlier samples from the Northern Mountains are labeled.

Table 3-5. Eigenvectors of metrics in the PCA.
Only metrics with eigenvectors greater than 0.2 on either of the first two axes are shown. Not shown are metrics that are redundant with those shown (percent dominant, percent 2 dominant, percent 3 dominant, etc.). The metric "% Diptera" was transformed by the arcsin square root function. Other transformed metrics were not retained in this table.

	Eigenvectors	
Metric	**1st axis**	**2nd axis**
Total taxa	**-0.24**	**0.21**
EPT taxa	**-0.25**	0.14
Plecoptera taxa	**-0.21**	0.10
Trichoptera taxa	**-0.21**	0.14
Diptera taxa	-0.11	**0.20**
% EPT	-0.10	**-0.30**
% Ephemeroptera	-0.03	**-0.26**
% Diptera	0.10	**0.31**
HBI	0.15	**0.28**
Intolerant taxa	**-0.25**	0.12
% 5 dominant	**0.27**	-0.01
Shannon-Wiener	**-0.27**	0.00
% Scrapers	-0.05	**-0.32**
% Collectors	0.09	**0.26**
Clinger taxa	**-0.23**	0.14
% Clingers	-0.05	**-0.23**

Table 3-6. Sample sizes in the bioregions of Idaho by *a priori* stream condition (least impacted or stressed) and data set assignment (calibration or test).

Ecoregion	**Least Impacted Calibration/Test**	**Stressed Calibration/Test**
Northern Mountains	17 / 9	13 / 7
Central and Southern Mountains	54 / 26	22 / 12
Basins	29 / 15	58 / 33
Totals	100 / 50	93 / 52

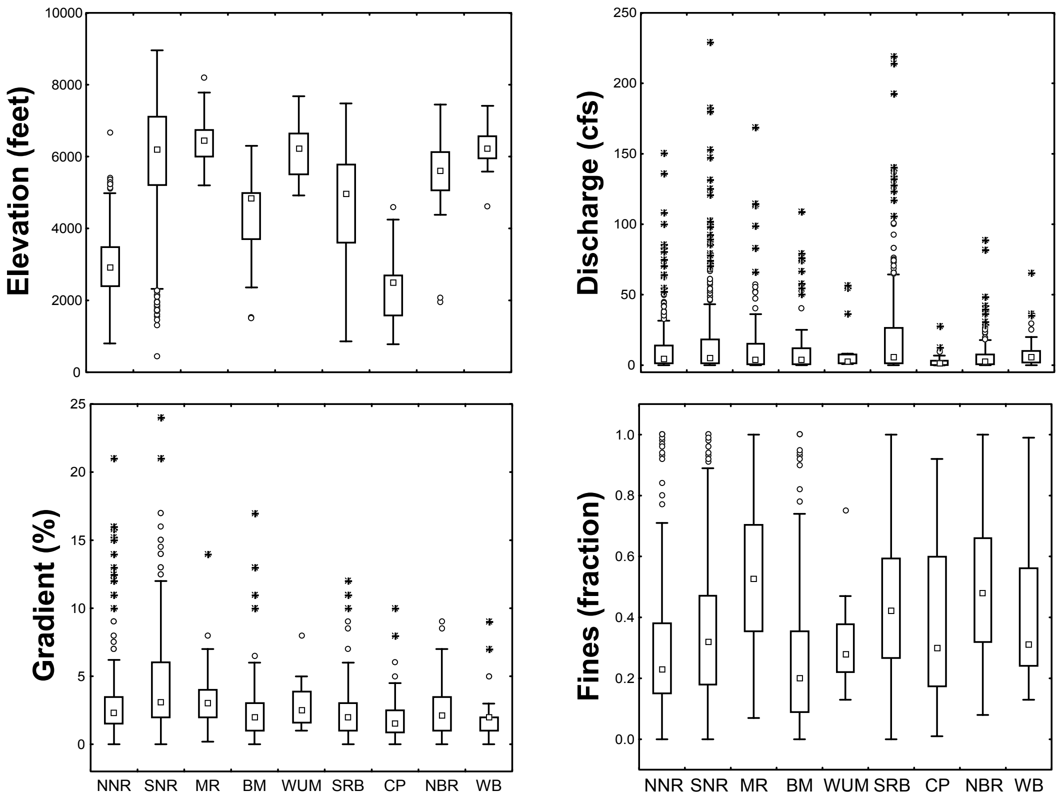


Figure 3-6. Physical characteristics of all sampled sites in the ecoregions of Idaho. The Northern Rockies are separated into northern (NNR) and southern (SNR) sub-ecoregions.

Index Metric Selection

Thirty-six biological metrics in six metric categories were calculated from the benthic macroinvertebrate taxa abundances (see Table 3-2). Taxa were identified to the lowest practical taxonomic level, usually genus or species. Laboratory protocol called for identification of 500 organisms from each composited sample; 37 percent of the samples were within 20 percent of this target. Three Hess sample grabs may fail to capture 500 or more organisms because of low densities caused by either oligotrophic or toxic conditions (natural or unnatural stressors). Very high densities may likewise be associated with conditions favorable to pollution intolerant or tolerant organisms. Taxa richness is expected to increase as more individuals are sampled, though composition metrics are independent of sample size. Total taxa richness was compared to subsample size in least impacted and stressed sites (Figure 3-7). At subsample sizes greater than 600, the least impacted taxa richness continued to increase with subsample size. Stressed taxa richness did not increase greatly with increasing subsample size. Though variability in subsample size was initially a concern, data were not re-sampled or eliminated from the analysis for three reasons. Richness in large subsamples increased only in least impacted sites, low densities may be a natural condition that should not be disregarded, and mathematical re-sampling (rarefaction; Hulbert 1971) would make routine application of the index more difficult.

Table 3-7. Median scores of assessed habitat features in the ecoregions of Idaho (scores from least impacted sites only).

	NNR	SNR	MR	BM	WUM	SRB	CP	NBR	WB
Substrate	12	10	6	12	9	10	13	8	10
Instream cover (fish)	15	17	16	17	17	15	14	15	17
Embeddedness	13	16	14	17	16	13	12	13	15
Velocity/Depth	15	15	15	10	15	15	13	15	9
Channel Shape	5	10	10	11	8	6	6	6	7
Pool/Riffle Ratio	4	2	3	1	1	4	4	2	0
Width/Depth Ratio	5	7	8	7	6	6	2	7	9
Bank Vegetation	7	8	7	10	10	8	4	9	10
Bank Stability	8	9	10	10	9	8	10	5	10
Disruptive Pressures	8	9	8	10	9	8	4	6	9
Zone of Influence	8	8	7	9	8	6	4	5	6
Total Score	93	112	104	110	112	103	89	89	107

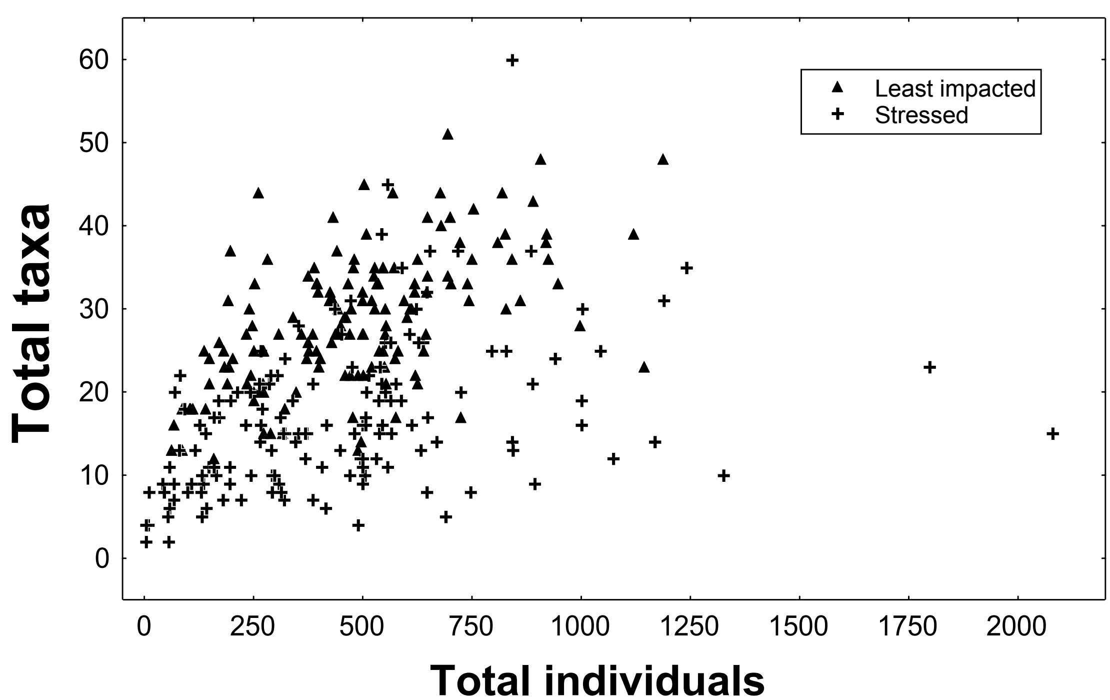


Figure 3-7. Comparison of total taxa and subsample size. The target substample size was 500 organisms.

During metric testing and initial index development, only calibration data was used. The discrimination efficiencies (DE) of the individual metrics were calculated as the percentage of stressed samples with metric values worse than the worst quartile of the least impacted metric values. Discrimination efficiencies of the 36 metrics are shown in Table 3-8. In all metric categories except voltinism, at least one metric had a DE of 75 percent or greater. Richness metrics performed well except for the Diptera taxa. Of the composition metrics, the percent Plecoptera metric had the best DE, while percent EPT also performed well. Intolerant taxa and HBI were better indicators of community pollution tolerance than was the percent tolerant metric. Of the diversity metrics, the two indexes and percent dominance using three or more taxa performed well. Scraper taxa richness and clinger taxa richness performed best in the feeding group and habit metric categories, respectively. The distributions of selected highly discriminating metrics are shown in Figure 3-8 and other metrics are included in Appendix A.

The lowest discrimination efficiencies were observed in the Northern Mountains. The most responsive metrics in this region were those related to overall taxa richness, total taxa, dominance, and diversity (Table 3-8). This suggests that the dominant stressor in the region may be predominantly those with moderate impact, such as logging operations. Because of their relative remoteness, the Northern Mountains may be less intensively used, resulting in stressed sites that are not as stressed compared to sites in accessible and populated regions.

Table 3-8. Discrimination efficiencies (DE) of candidate metrics in three bioregions, statewide DE, and metric trend with increasing impairment (▼ – decreasing, △ – increasing). Calibration data only.

	No. Mtns.	C&S Mtns	Basins	Idaho - ALL	Trend
Richness metrics					
Total taxa	64	71	78	74	▼
EPT[1] taxa	50	86	93	85	▼
Ephemeroptera taxa	50	81	91	83	▼
Plecoptera taxa	43	91	85	80	▼
Trichoptera taxa	29	71	86	74	▼
Diptera taxa	21	38	35	33	▼
Composition metrics					
Percent EPT	36	81	81	74	▼
Percent Ephemeroptera	36	71	66	62	▼
Percent Plecoptera	36	86	86	79	▼
Percent Trichoptera	50	67	71	67	▼
Percent Hydropsychidae					
Percent Elmidae	43	33	69	57	▼
Percent Diptera	14	62	47	45	△
Percent Diptera (non-chironomid)	7	48	31	31	△
Percent Chironomidae	14	57	55	50	△
Percent Non-insects	14	33	78	58	△
Pollution tolerance metrics					
Intolerant taxa (0-1)	36	91	97	86	▼
Intolerant taxa (0-3)	50	91	91	85	▼
Percent Tolerant	21	52	69	58	△
HBI	21	81	85	74	△
Diversity metrics					
Percent Dominant	50	52	74	66	△
Percent two dominant	64	71	83	77	△
Percent three dominant	71	91	90	87	△
Percent five dominant	71	91	90	87	△
Percent 10 dominant	79	95	88	88	△
Shannon-Wiener	57	81	88	82	▼
Simpson's	64	71	85	79	
Feeding group metrics					
Scraper taxa	29	81	85	75	▼
Percent Scraper	36	71	85	74	▼
Percent Predator	14	67	66	58	▼
Percent Collector	36	43	60	53	△
Habit metrics					
Clinger taxa	43	76	93	82	▼
Percent Clinger	29	52	81	67	▼
Percent Clingers of insects	29	48	74	61	▼
Voltinism metrics					
Semi-voltine taxa	14	24	53	41	
Percent Semi-voltine	36	29	64	52	

[1]EPT: Ephemeroptera, Plectoptera, Trichoptera
[2]HBI = Hilsenhoff Biotic Index

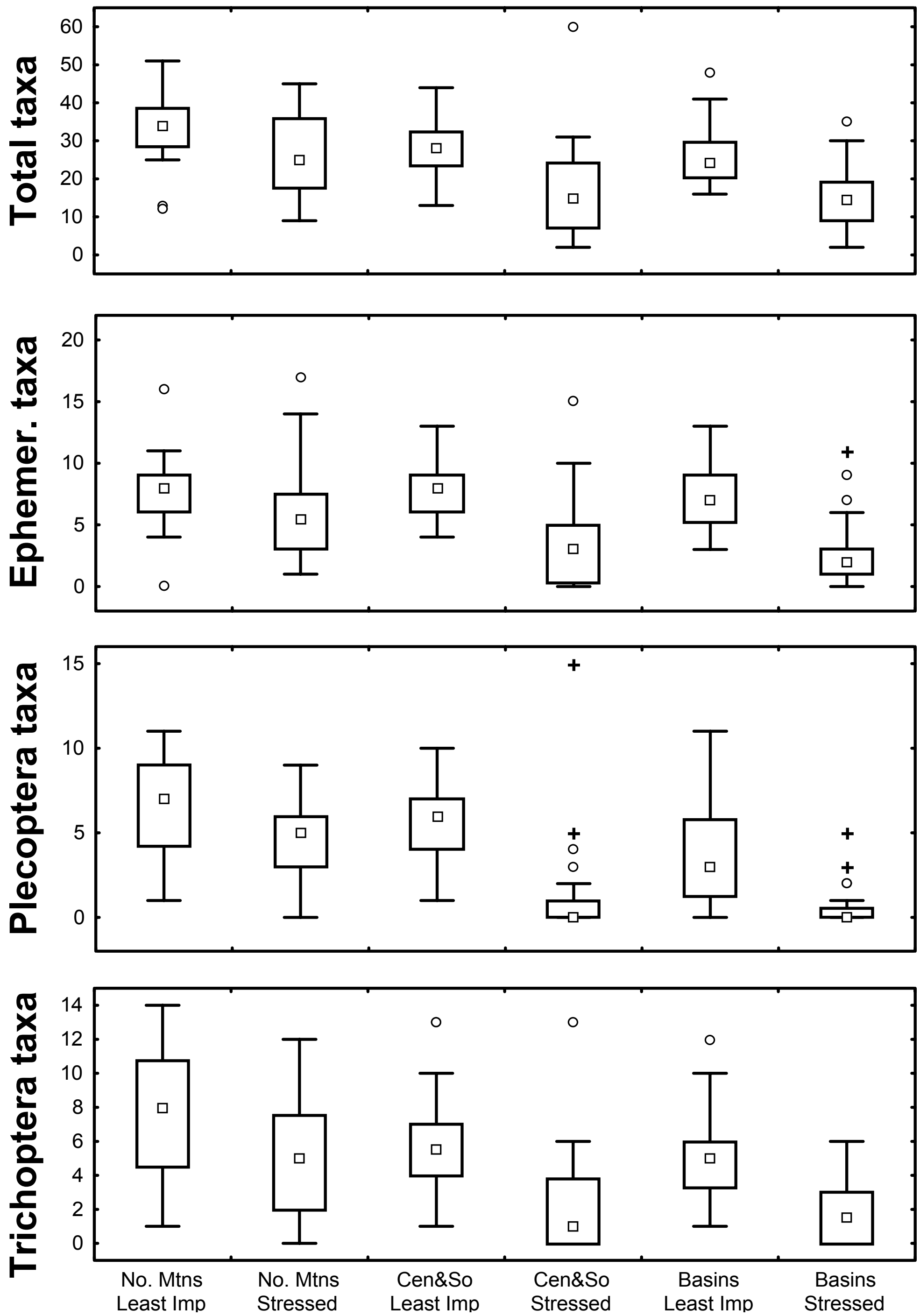


Figure 3-8. Selected distributions of least impacted and stressed metrics in bioregions of Idaho; calibration samples only.
Refer to legend in Figure 3-2.

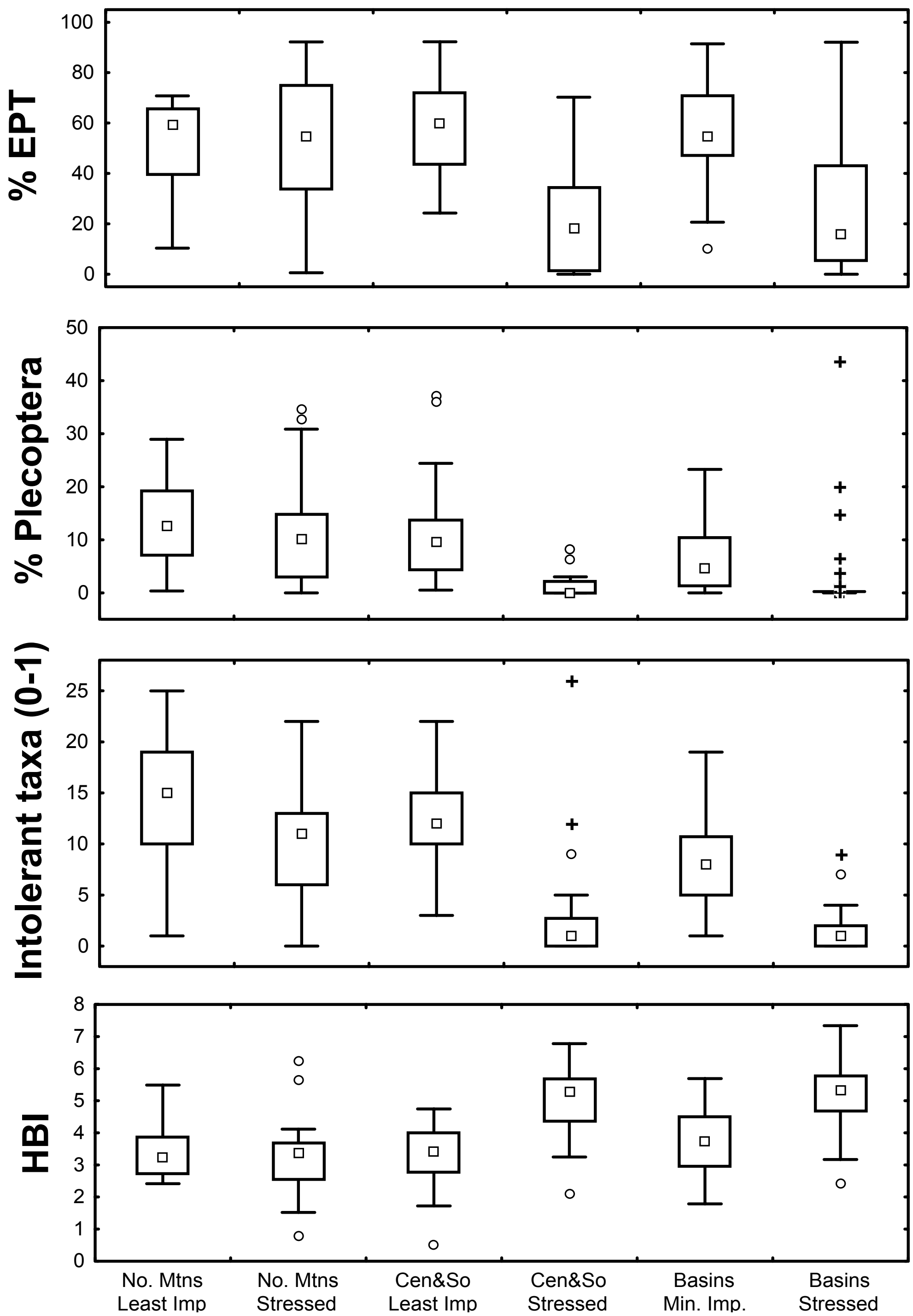


Figure 3-8. *(continued)*

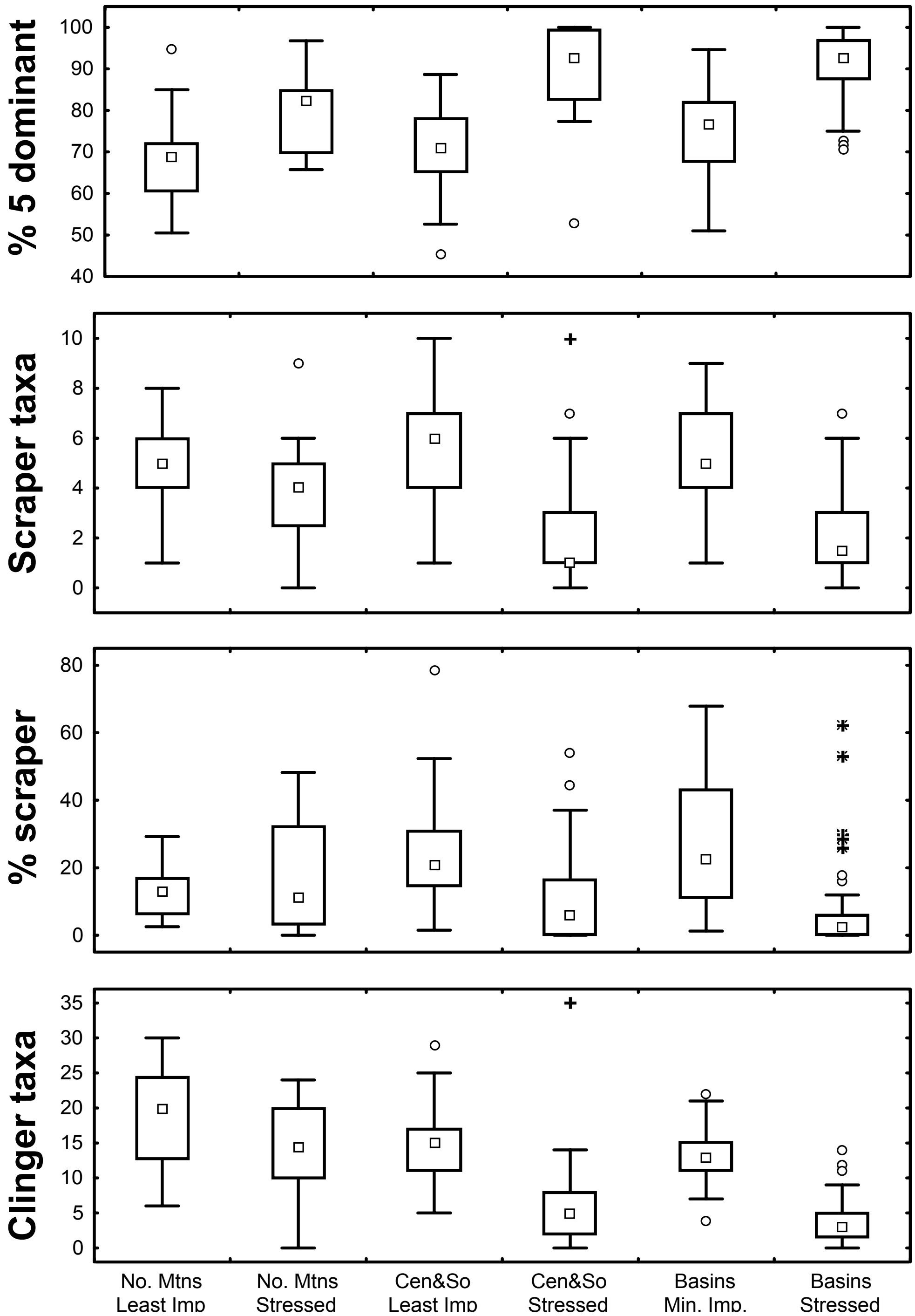


Figure 3-8. *(continued)*

Index Development

Alternative indexes were calculated as the average of metric scores. Metrics were included and interchanged in the index until the combination which best satisfied the index development goals (to include responsive metrics from all of the metric categories, to include metrics that are not redundant, and to show the greatest separation between least impacted and stressed index scores [expressed as the index DE]) was identified.

Fourteen metrics were selected as candidates for inclusion in the index based on robust discrimination between least impacted and stressed sites in the three bioregions. All metric categories were represented except voltinism. Redundancy between the metrics was checked using correlation analysis, with metrics considered excessively redundant if the correlation coefficient was greater than or equal to 0.90 (Table 3-9).

All multimetric index alternatives included four taxa richness metrics: total taxa, Ephemeroptera taxa, Plecoptera taxa, and Trichoptera taxa. The EPT taxa metric performed better than any of its components individually, but was redundant with total taxa and was eliminated in favor of the broader measure of all taxa. The inclusion of the three component metrics in the index allows the important insect orders to convey more distinct information into the assessment. The Plecoptera are represented as taxa counts and relative abundance in the index alternatives. The two metrics are not highly correlated and each conveys unique information about diversity and predominance of this sensitive taxa group. The HBI and intolerant taxa (0 to 1) were interchanged while testing alternatives. The intolerant taxa (0 to 3) metric was not included because of redundancy with total taxa and clinger taxa. Though the number of intolerant taxa outperformed the abundance weighted HBI in metric DE analysis, the index DE was higher when it included the HBI. Three dominance metrics, using three, five, and 10 taxa, were interchanged in some alternatives.

The index that best satisfied the three index development goals had a DE of 88.2 percent. The Stream Macroinvertebrate Index (SMI) includes the following nine metrics:

Total taxa	Trichoptera taxa	Percent 5 dominant taxa
Ephemeroptera taxa	Percent Plecoptera	Scraper taxa
Plecoptera taxa	HBI	Clinger taxa

Table 3-9. Correlation coefficients (Pearson product-moment r) for metrics that show good discrimination efficiency. Values greater than 0.85 or less than -0.85 are bold-typed.

	Total taxa	EPT taxa	Ephemeroptera taxa	Plecoptera taxa	Trichoptera taxa	% EPT	% Plecoptera	% Trichoptera	HBI	Intolerant taxa (0-1)	Intolerant taxa (0-3)	% 3 dominant	% 5 dominant	% 10 dominant	Shannon-Wiener	Scraper taxa	% Scrapers	Clinger taxa
Total taxa																		
EPT[1] taxa	**.92**																	
Ephemerop. taxa	.80	**.87**																
Plecoptera taxa	.77	**.86**	.63															
Trichoptera taxa	.82	**.86**	.60	.63														
% EPT	.26	.43	.40	.40	.31													
% Plecoptera	.26	.38	.23	.53	.24	.45												
% Trichoptera	.20	.23	.09	.13	.37	.31	.01											
HBI[2]	-.38	-.56	-.48	-.54	-.44	-.77	-.53	-.37										
Intolerant (0-1)	.84	**.94**	.81	.83	.82	.41	.37	.20	-.59									
Intolerant (0-3)	**.89**	**.97**	.82	**.88**	.83	.41	.40	.20	-.59	**.97**								
% 3 dominant	-.69	-.68	-.59	-.60	-.58	-.47	-.41	-.26	.59	-.64	-.67							
% 5 dominant	-.74	-.74	-.64	-.65	-.62	-.44	-.40	-.25	.57	-.69	-.72	**.95**						
% 10 dominant	-.82	-.80	-.70	-.70	-.69	-.35	-.36	-.23	.49	-.75	-.78	.83	**.91**					
Shannon-Wiener	.75	.74	.65	.64	.63	.48	.41	.25	-.60	.69	.72	**-.95**	**-.93**	-.84				
Scraper taxa	.72	.75	.73	.52	.69	.34	.12	.23	-.43	.71	.71	-.54	-.58	-.60	.58			
% Scrapers	.16	.26	.26	.18	.22	.41	.04	.20	-.49	.27	.24	-.29	-.26	-.19	.29	.44		
Clinger taxa	**.89**	**.93**	.79	.77	.84	.37	.28	.24	-.51	**.88**	**.91**	-.65	-.70	-.76	.71	.78	.27	

[1]EPT = Ephemeroptera, Plectoptera, Trichoptera
[2]HBI = Hilsenhoff Biotic Index

Table 3-9. Correlation coefficients for metrics that show good discrimination efficiency.

Of the proposed index metrics, the strongest correlation was between total taxa and clinger taxa (r = 0.89). Though the correlation coefficient was high, these metrics were both retained in the index because they represented important information from different metric categories. Correlations between other metrics were not as strong (r < 0.85). The seemingly redundant metrics Plecoptera taxa and percent Plecoptera had a correlation coefficient of only 0.53, demonstrating the difference in richness and composition measures of the assemblage. These two were not redundant because a single Plecopteran species could be abundant, resulting in a high percent Plecoptera. The degree of separation between least impacted and impaired site index values can be seen in box and whisker diagrams of index score distributions in the bioregions (Figure 3-9).

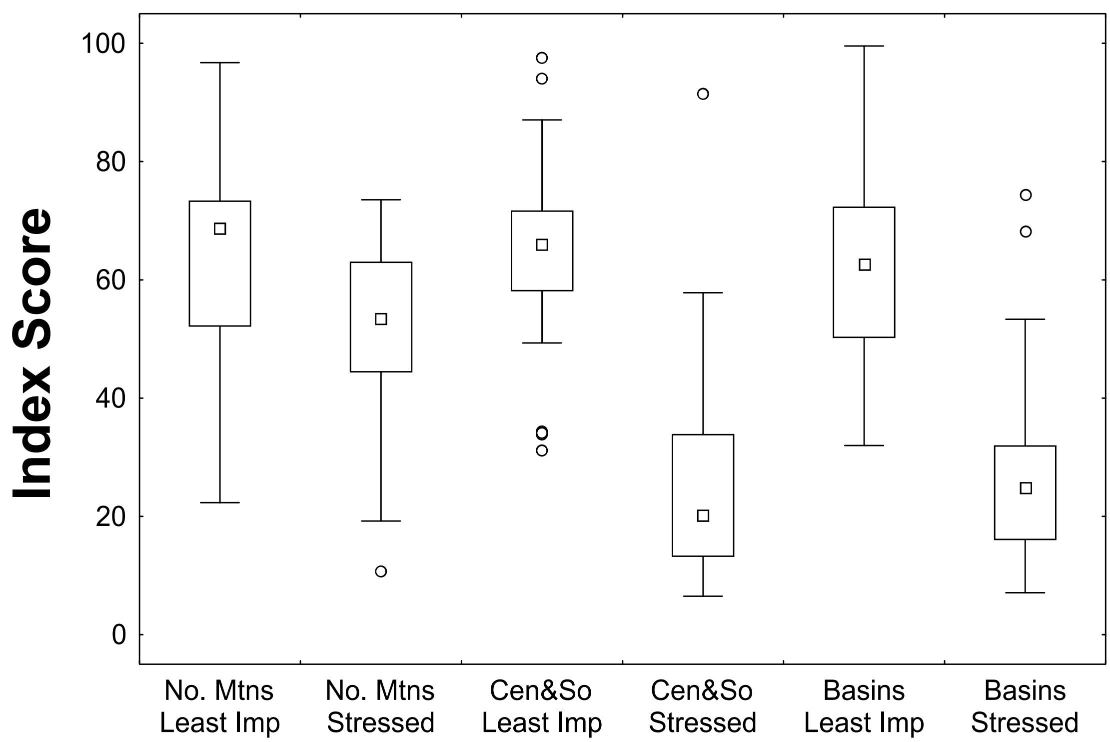


Figure 3-9. Index score distributions in least impacted and stressed sites, index development data.

The 25th percentiles of least impacted scores (the lower edge of the boxes) were used to judge discrimination efficiency.

The greatest degree of overlap between least impacted and stressed index scores is seen in the Northern Mountains. Though index alternatives that used the best performing metrics in the Northern Mountains were evaluated, none out-performed the proposed index. This may be attributed to relatively moderate stressors in this sparsely developed bioregion.

Index Tested with Independent Data and Compared to Other Indexes

The SMI was tested using the 25th percentile of least impacted calibration data as a threshold to which the randomly designated test data was compared. The index showed greater discriminating ability with test data than it did with calibration data (Table 3-10, Figure 3-10). The proposed index appears to be robust.

Table 3-10. Discrimination efficiencies for the SMI using independent test data.

	Northern Mountains	Central and Southern Mountains	Basins	Idaho - all
Least impacted	88.9	80.8	80	82.0
Stressed	71.4	100	93.9	92.3

The SMI was compared to several other indexes that have been applied in Idaho. For each index, discrimination abilities were calculated using all data (calibration and test). In comparison to the other indexes, the proposed index performed up to 10 percent better. The River Macroinvertebrate Index (Royer et al. 2001) was developed for use in larger rivers of Idaho and did not perform as well as the proposed index in any of the three bioregions (see Table 3-11). The data set under investigation is predominantly small wadeable streams, and the river index may be better suited to larger streams and rivers only.

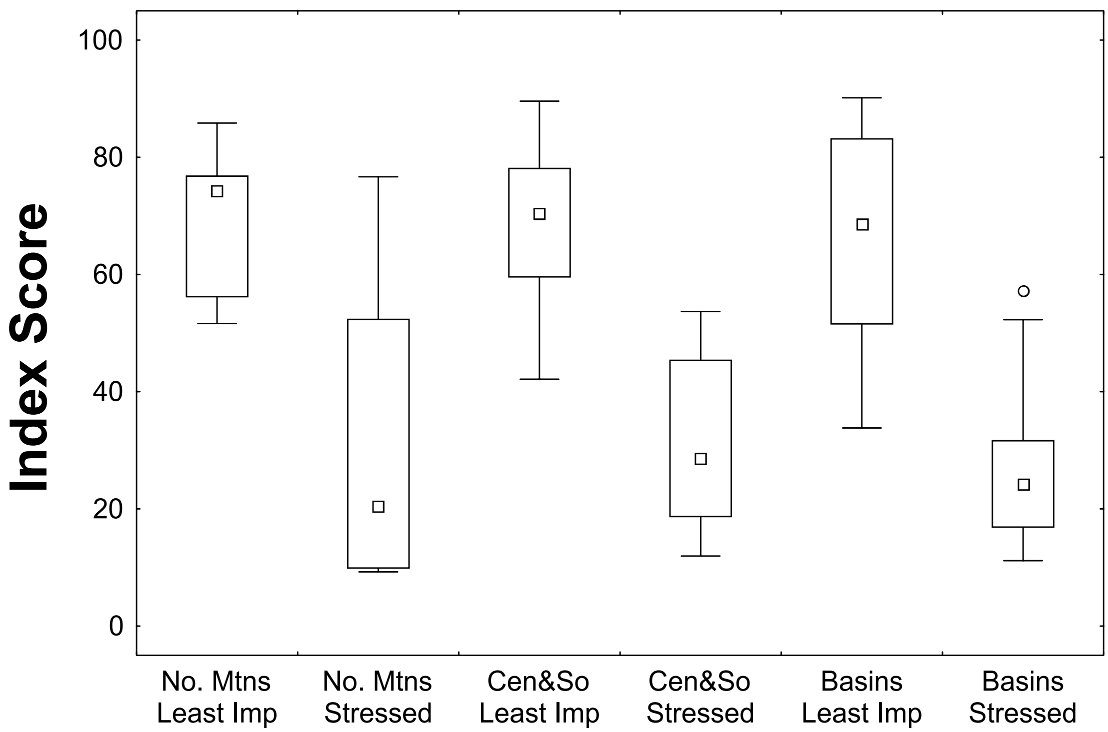


Figure 3-10. Index score distributions in least impacted and stressed sites, test data. The 25th percentiles of least impacted scores (calibration data, cf. Figure 3-9 were used to judge discrimination efficiency).

The Macroinvertebrate Biotic Index (described in Mebane 2000) developed by DEQ staff for statewide application performed equal to the SMI in the Basins bioregion (see Table 3-11), but identified 10 percent fewer impaired sites in the two mountain bioregions. It includes both total taxa and EPT taxa, that are redundant by our standards (Pearson product-moment correlation coefficient > 0.9).

The Benthic Index of Biotic Integrity (Kerans and Karr 1994) performed identically to the SMI in the Basins and the Northern Mountains but performed less efficiently (9 percent or 3 sites) in the Central and Southern Mountains (see Table 3-11). The two indexes have six metrics in common, but the Benthic Index of Biotic Integrity includes some metrics that did not show strong discrimination ability with the current data (percent predators and semi-voltine taxa).

The ecoregion-specific indexes developed by Robinson and Minshall (1998) were compared to the SMI within three ecoregions only (Northern Rockies, Snake River Plain, and Northern Basin and Range; Table 3-12). Roughly half of the metrics are composition metrics at the genus and family taxonomic levels. Discrimination efficiency of the Northern Basin and Range index was 100 percent, but the Snake River Plain and Northern Rockies were 84 percent and 68 percent, respectively (Table 3-12). Though the Robinson and Minshall indexes performed well, they are not applicable in all regions of the state.

Table 3-11. SMI configuration (Index A) and other indexes used in Idaho. Indexes are composed of the metrics opposite the index symbol. Discrimination efficiencies (DE, %) are based on the 25th percentile of least impacted index scores within the bioregions defined by this study.

	SMI	**RMI**[1]	**MBI**[2]	**B-IBI**[3]
Total taxa	A	RMI	MBI	B-IBI
EPT taxa		RMI	MBI	
Ephemerop. Taxa	A			B-IBI
Plecoptera taxa	A			B-IBI
Trichoptera taxa	A			B-IBI
Percent EPT			MBI	
Percent Plecoptera	A			
Percent Elmidae		RMI		
HBI	A		MBI	
Intolerant (0-1)				B-IBI
Percent Tolerant				B-IBI
Percent Dominant (# taxa)	A (5)	RMI (1)	MBI (1)	B-IBI (3)
Shannon-Wiener			MBI	
Scraper taxa	A			
Percent Scraper			MBI	
Percent Predators		RMI		B-IBI
Clinger taxa	A			B-IBI
Semi-voltine taxa				B-IBI
Northern Mountains (stressed n = 21)	57%	52%	48%	57%
Central and Southern (stressed n = 33)	97%	85%	85%	88%
Basins (stressed n = 91)	95%	91%	95%	95%
All Bioregions	89%	84%	85%	87%

[1]River Macroinvertebrate Index (Royer et al. 2001)
[2]Macroinvertebrate Biotic Index (described in Mebane 2001)
[3]Benthic Index of Biotic Integrity (Kerans and Karr 1994)

Table 3-12. SMI configuration (Index A) and other indexes calibrated for the Northern Rockies, Snake River Plain, and Northern Basin and Range of Idaho (Robinson and Minshall 1998).
Indexes are composed of the metrics opposite the index symbol. Discrimination efficiencies (DE, %) are based on the 25th percentile of least impacted index scores within the ecoregions. Metric calibration for SMI was based on bioregional statistics, but only sites within the three ecoregions were included for this comparison.

	SMI	**NR-I**	**SRP-I**	**NBR-I**
Total taxa	A			
EPT taxa		NR-I	SRP-I	NBR-I
Ephemeroptera taxa	A			
Plecoptera taxa	A			
Trichoptera taxa	A			
Percent EPT		NR-I	SRP-I	NBR-I
Percent Plecoptera	A			
Percent Chironomidae				NBR-I
EPT/Chironomidae		NR-I		
EPT/Chiron. + Oligochaeta			SRP-I	
HBI	A	NR-I	SRP-I	
Percent dominant (# taxa)	A (5)	NR-I (1)	SRP-I (1)	
Simpson's Index		NR-I		
Scraper taxa	A			
Percent scraper		NR-I		NBR-I
Percent filterers			SRP-I	NBR-I
Percent shredders		NR-I		
Clinger taxa	A			
Percent Baetis		NR-I		
Percent Heptageniidae				NBR-I
Percent Zapada		NR-I		NBR-I
Percent Hydropsychidae		NR-I		NBR-I
Percent Rhyacophilidae			SRP-I	
Percent Brachycentrus		NR-I	SRP-I	
Percent Ephemerellidae				NBR-I
Percent Drunella		NR-I		
Percent Capniidae			SRP-I	
Percent Elmidae		NR-I	SRP-I	NBR-I
Percent Simulidae		NR-I	SRP-I	
Percent Turbellaria			SRP-I	
Northern Rockies (stressed n = 37)	76%	68%		
Snake River Plain (stressed n = 56)	93%		84%	
No. Basin & Range (stressed n = 24)	96%			100.0%
Three Ecoregions	88%		82%	

Northern Mountains Recalibration and Index Finalization

After the initial development of the SMI and before establishing the final metric scoring criteria, the data set and metric calculations were refined in two ways. First, least impacted and stressed sites were redefined in the Northern Mountains. The data collected between 1994 and 1996 was augmented with data collected through 1999. The new data set included 31 least impacted sites and 38 stressed sites that better representated the best and worst conditions compared to the earlier data set (see Appendix D). The second refinement was a reexamination of the calculation of richness metrics. Most taxa groups were identified at a consistent taxonomic level (e.g., genus). However, some individuals of a sample were identified at family level or higher because the specimens lacked distintive characteristics (damaged or immature specimens). The uniqueness of these individuals identified at higher levels was ambiguous and counting them as unique may artificially increase measures of community diversity. Whereas preliminary calculations assumed they were unique, the final assumption was that they were not unique in samples where lower level specimens were also identified.

Refinement of the data set significantly improved the performance of the SMI in the Northern Mountains (Figure 3-11). Specifically, the index discrimination efficiency (DE) increased from 62 percent to 90 percent using the 25th percentile of least impacted sites as the criterion (Table 3-13). The DE in all bioregions was 94 percent also using the 25th percentile. Almost 75 percent of stressed site index scores were below the *minimum* least impacted scores. Scoring formulas for the final index metrics are presented in Appendix B. Five sites had index scores that appeared as outliers (Table 3-14). Outliers may signify natural variability, misclassification of stream conditions *a priori*, or an under-represented site class.

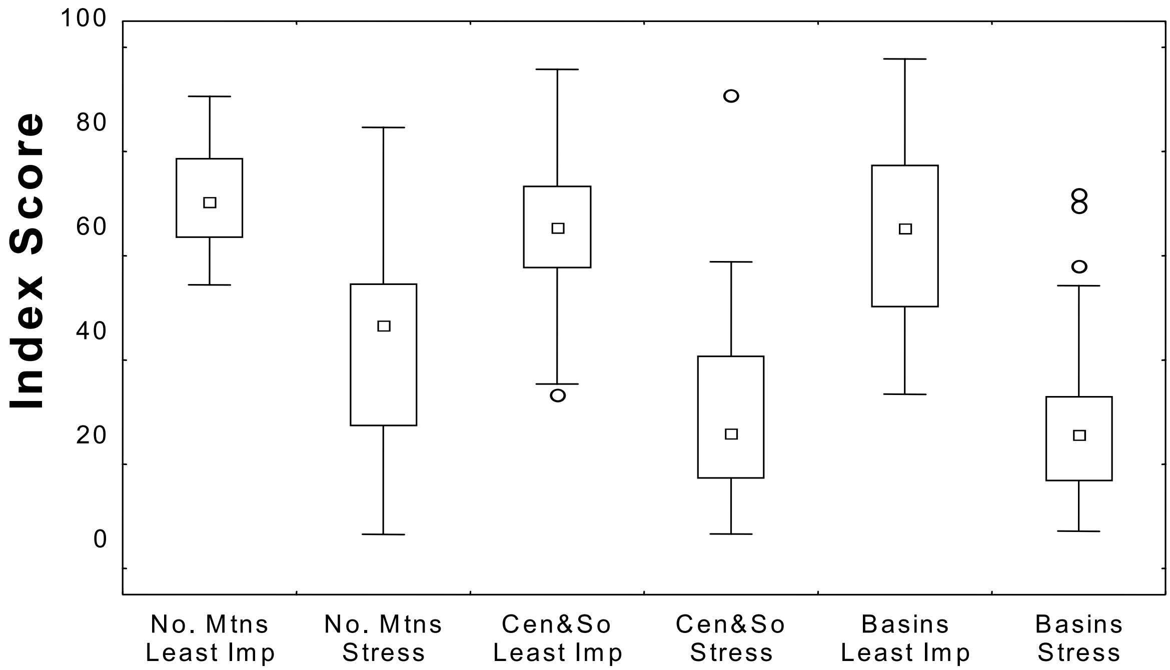


Figure 3-11. SMI score distributions in least impacted and stressed sites, calibration and test sites combined.

The SMI was recalibrated using updated least impacted and stressed sites in the Northern Mountains and eliminating ambiguous taxa in the other bioregions.

Table 3-13. Statistics of SMI. Least impacted distributions and DE's based on various percentiles.

	(n)	Min.	5th	10th	25th	median	75th	90th	Max.
Northern Mountains									
Scores (Least impacted)	(31)	39*	55	56	64	70	79	82	91
DE (% of stressed)	(38)	74%	79%	79%	90%	92%			
Central and Southern Mountains									
Scores (Least impacted)	(80)	33	36	50	58	65	73	81	96
DE (% of stressed)	(33)	70%	70%	91%	94%	97%			
Basins									
Scores (Least impacted)	(44)	33	34	42	50	64	77	88	98
DE (% of stressed)	(91)	75%	76%	89%	96%	98%			

* Minimum reference score of all Northern Mountains sites.

Table 3-14. Samples with outlying index scores.

Bioregion	Site ID	Stream name	Ecoregion	Condition	Index Score
C&S Mtns.	95EIRO0A92	Lower Yankee Fork	So. Northern Rockies	Stressed	90.6
C&S Mtns.	95SEIRO061	Beaver Creek	Wasatch & Uinta Mtns.	Least Imp.	33.2
Basins	96SCIROA07	Shoshone Creek	Snake River Basin	Stressed	69.4
Basins	96SCIROB68	Lower Chimney Creek	Snake River Basin	Stressed	71.7
Basins	96SCIROB18	Lower Big Creek	Snake River Basin	Stressed	57.9

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Of the three project objectives, two were accomplished through the above analysis and the third (to recommend appropriate applications of the index) is completed in the "Index Rating and Application" section below. A regionally-calibrated multimetric biological index for Idaho streams using the benthic macroinvertebrate assemblage was successfully developed and favorably evaluated. The Stream Macroinvertebrate Index (SMI) is robust and repeatable in the three bioregions of Idaho. The index incorporates nine benthic macroinvertebrate metrics from six metric categories, combining a range of ecological information on biological conditions in wadeable streams. The index "correctly" classifies 94 percent of the stressed sites below the 25^{th} percentile of least impacted scores, showing excellent agreement with the stream condition assigned *a priori* by DEQ personnel.

The ecological significance of the nine metrics in the SMI can be described individually to illustrate the range of environmental influences on the index values. There are four richness metrics: total taxa, Ephemeroptera taxa, Plecoptera taxa, and Trichoptera taxa. Total taxa richness is a primary measure of biodiversity. Diversity of the benthic macroinvertebrate community indicates heterogeneity of habitat, complexity of trophic interactions and community dynamics, and a wide range of food resources. Diversity within each of the three insect orders confers similar information, but each of these orders is generally understood to have several species that are sensitive to pollution. High diversity within each of the orders therefore implies that favorable habitats, community dynamics, and food resources exist in the absence of pollution. The composition metric in the index, "Percent Plecoptera", confers information regarding the degree to which the stream can support larger populations of stoneflies. Stoneflies require clean, cold, and well oxygenated water for survival; conditions that are adversely impacted by a wide range of human activity. The HBI calculates the predominance of pollution tolerant individuals of all taxa groups, and thereby describes the likely existence of pollutants or perturbations at a site.

Predominance in the macroinvertebrate community by the top five taxa might indicate that conditions are only suitable for a limited type of organism. Such a condition may arise from limited food resource types, habitat degradation, or competitive displacement of native taxa by exotic species. High diversity in scraper taxa probably means conditions are suitable for a productive periphyton community—conditions including sufficient nutrients, non-toxic water, and stable substrate. A high diversity of clinger taxa likewise indicates stable substrate, as well as faster currents, as are expected in undisturbed stream channels of Idaho.

An understanding of the metric components of the index will aid in the appropriate application of the index. The index is a composite of the biological information, not necessarily a diagnostic tool, but a general status indicator. If a single metric value is suboptimal but all others in the index resemble least impacted values, it would be difficult to discern impairment from the index. However, if several metrics are suboptimal, that would be reflected in the index value and further investigation might identify the source or sources of stress to the community. To apply the biological information for diagnostic purposes would require examination of individual metric responses to identify specific stressors and calibration of those stress-response relationships.

Index Rating and Application

With the response to stressors shown by the SMI in Figure 3-11, it is possible to propose a rating system for the index. The multimetric index value for a site is a summation of the scores of the metrics and has a finite range within each stream class and index period depending on the maximum possible scores of the metrics (Barbour et al. 1999). This range can be subdivided into any number of categories corresponding to various levels of impairment. Because the metrics are normalized to least impacted conditions and expectations for the stream classes, any decision on subdivision should reflect the distribution of the scores for the least impacted sites.

The number of rating categories that can be discerned is determined by the variability of the index in the least impacted sites. The interquartile ranges of the least impacted site scores are 15 to 27 points in the three bioregions (see Table 3-13). This suggests that five categories can be reliably discerned on the 100 point index scale (Fore et al. 1994, 1996).

In a suggested rating scheme, the ratings of "very good" or "good" are applied to sites with scores greater than the 25th percentile of the calibration least impacted index scores (Table 3-15). The threshold between "very good" and "good" is the mid-point between the 25th percentile and the maximum index score (100). Ratings of "fair," "poor," or "very poor" are assigned to sites with scores below the 25th percentile. The subdivisions below the 25th percentile are based on a trisection of the scale to the minimum index score (0). This or any alternative rating system allows rapid prioritization of sites by biological condition. Index discrimination efficiencies based on the minimum, 10th, 25th, and 50th percentiles are presented in Table 3-13. Adjustments to category thresholds or narrative labels based on the least impacted distributions will allow avoidance of Type I or Type II errors as needed. The recommended rating scheme consistently allows a Type I error of 25 percent, because the 25th percentile of least impacted sites is used as the threshold between "good" and "fair" ratings.

Using a percentile of the reference scores (e.g., the 25th) as a threshold results in that percentage of reference sites failing the criterion. Impairment, as measured by the index score, is a gradual continuum of condition. Yet, in a management context, a threshold value is required to trigger management action. The choice of the threshold reflects a tolerance for risk and uncertainty. Risks include the risk of declaring a good site impaired (false positive; statistical Type I error) and the risk of declaring an impaired site good (false negative; statistical Type II error). Assuming either risk is accepting error; there may be political consequences associated with not protecting or over-protecting Idaho's aquatic resources. An additional source of uncertainty comes from the selection of reference sites that truly represent least-stressed conditions.

A 25th percentile of least impacted conditions is commonly chosen because it is deemed sufficiently conservative to protect aquatic resources, and reflects some uncertainty in the reference site selection. Some states have selected a 10th percentile threshold if they have greater confidence that their reference sites are not stressed and their methods yield precise results (e.g., Maryland, West Virginia). Nevertheless, a lower percentile reduces the power to detect impairment. Furthermore, a decline in stream condition from the median to the 10th percentile (e.g., a decline of an index score from 64 to 42 in the Basins) would not trigger a management response, though this change in condition may be unacceptable in the context of antidegradation policy.

Table 3-15. Example rating categories based on 25th percentiles of least impacted SMI scores.

Rating	Northern Mountains	Central and Southern Mountains	Basins
	Index score range		
Very Good (midpoint between 25th percentile and maximum index score to maximum score)	84 - 100	80 - 100	76 - 100
Good (25th percentile to midpoint between 25th percentile and maximum score)	65 - 83	59 - 79	51 - 75
Fair (upper trisect of minimum score to 25th percentile)	44 - 64	40 - 58	34 - 50
Poor (middle trisect of minimum score to 25th percentile)	22 - 43	20 - 39	17 - 33
Very Poor (lower trisect of minimum score to 25th percentile)	0 - 21	0 - 19	0 - 16

Application of the biological SMI and rating system in Idaho could proceed as follows for any new biomonitoring sites:

(1) Collect biological sample and associated data.
(2) Identify organisms in a subsample to standard taxonomic level.
(3) Calculate the nine index metrics.
(4) Score index metrics using formulas (Appendix B).
(5) Calculate the biological SMI as an average of the nine metric scores.
(6) Rate the site's biological condition (Table 3-15 or alternative).
(7) Interpret rating in context of water resource management decisions.

Recommendations

- The multimetric macroinvertebrate index proposed in this study includes nine metrics: total taxa, Ephemeroptera taxa, Plecoptera taxa, Trichoptera taxa, percent Plecoptera, HBI, percent five dominant taxa, scraper taxa, and clinger taxa. The index should be applied in the bioregions of Idaho in its current initial form, with the understanding that improvements may be forthcoming as new data are incorporated into future analyses.

- The index developed here is appropriate for biocriteria, and a rating system can be determined for application of biocriteria. The use of a 25th percentile threshold automatically results in 25 percent of least impacted sites scoring "fair" or worse. The acceptance levels for Type I and Type II errors in identifying biological condition should be defined and threshold index values assigned accordingly. The rating system suggested above (25th percentile) is only one alternative.

- Improvements to the index should be pursued. Such improvements may include reclassification, redevelopment, or recalibration depending on the availability of new data. Redevelopment would include re-testing all metrics and new metric combinations. Recalibration involves readjustment of scoring thresholds of the initial index metrics. Index redevelopment should accompany large changes in data availability, site classification, or *a priori* stream condition status. Index recalibration should accompany less extreme changes in database characteristics.

- Because site classification was largely dependent on professional judgement (especially in under-represented ecoregions), effort should be made to collect additional data where it is sparse. The Columbia Plateau, Wyoming Basin and Wasatch-Uinta Mountains jointly constitute less than 10 percent of the land area in the state, but it is important to group these ecoregions into bioregions with greater confidence. Minimally impacted and stressed sites would complement future analyes.

- Subsampling procedures should be reviewed to identify possible ways to avoid extremely large subsample sizes.

- Some least impacted sites had low scores, rating in the "fair" or "poor" range using the suggested rating scheme. Extremely low scores and outliers among the least impacted sites may indicate misidentification of stream condition *a priori*. These outliers may not be representative of their bioregions (e.g., unique anomalous geology or water source). The outliers should be excluded from the least impacted set if there are clearly anomalous conditions at the sites that would make them non-representative, or if there is previously undetected anthropogenic stress or pollution. They should not be excluded from the least impacted set simply because of a low score.

- We did not examine precision of individual site scores in this study. Precision of scores can be estimated from repeated (replicate) observations at sites. Interannual variability can be estimated from repeat visits in different years. These components of variability can be used to develop expectations of the overall natural variability of least impacted sites and precision of the methods, to sharpen and improve the ability to detect impairment.

ACKNOWLEDGEMENTS

This biological index development and report grew out of earlier data analysis supported by EPA, Office of Water, Office of Science and Technology. Susan Jackson, EPA/ Office of Science and Technology , and Gretchen Hayslip, EPA Region 9, initiated the data analysis project and provided critical discussions and reviews. The Project Manager for the final analysis presented here was Michael Edmondson of DEQ. Bill Clark, Michael Edmondson, Cynthia Grafe, Mike McIntyre, Chris Mebane, and David Moser of DEQ reviewed results and earlier drafts, and provided critical ideas and discussion on classification of Idaho streams. The report was produced by Brenda Fowler and Jackie Bowersox.

REFERENCES

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of the North American Benthological Society. 15(2):185-211.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. Second Edition. EPA/841-B-99-002. U.S. EPA, Office of Water. Washington, D.C.

Barbour, M.T., J.L. Plafkin, B.P. Bradley, C.G. Graves, and R.W. Wisseman. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: metric redundancy and variability among reference stream sites. Environmental Toxicology and Chemistry. 11(4):437-449.

Barbour, M.T., J.B. Stribling, and J.R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Pages 63-77 in W.S. Davis and T.P. Simon (editors). Biological assessment and criteria. Tools for water resource planning and decision making. Lewis Publishers. Boca Raton, Florida.

Chandler, G.L., T.R. Maret, and D.W. Zaroban. 1993. Protocols for assessment of biotic integrity (fish) in Idaho streams. Water Quality Monitoring Protocol Number 6. Idaho Division of Environmental Quality. Boise, Idaho. 40 p.

Clark, W.H. and T.R. Maret. 1993. Protocols for assessment of biotic integrity (macroinvertebrates) for wadable Idaho Streams. Water Quality Monitoring Protocols Report 5, Idaho Division of Environmental Quality. Boise, Idaho.

Cummins, K.W. and M.J. Klug. 1979. Feeding ecology of stream invertebrates. Annual Review of Ecology and Systematics. 10: 147-172.

Davis, W.S. and T.P. Simon (editors). 1995. Biological assessment and criteria. Tools for water resource planning and decision making. Lewis Publishers. Boca Raton, Florida.

Fore, L.S., J.R. Karr, and L.L. Conquest. 1994. Statistical properties of an index of biological integrity used to evaluate water resources. Canadian Journal Fisheries and Aquatic Sciences. 51: 1077-1087.

Fore, L.S., J.R. Karr, and R.W. Wisseman. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of North American Benthological Society. 15: 212-231.

Gibson, G.R., M.T. Barbour, J.B. Stribling, J. Gerritsen, and J.R. Karr. 1996. Biological criteria: technical guidance for streams and small rivers. EPA 822-B-96-001. U.S. EPA, Office of Water. Washington, D.C.

Hayslip, G.A. (editor) 1993. Region 10 in-stream biological monitoring handbook for wadeable streams in the Pacific Northwest. EPA 910/9-92-013. U.S. EPA. Seattle, WA.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist. 20: 31-39.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family level biotic index. Journal of the North American Benthological Society. 7(1):65-68.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management. 10:629-635.

Division of Environmental Quality (DEQ). 1996. 1996 beneficial use reconnaissance project workplan. Idaho Division of Environmental Quality. Boise, Idaho.

Idaho Water Quality Standards and Wastewater Treatment Requirements. IDAPA 58.01.02. Office of Administrative Rules. Boise Idaho.

Jongman, R.H., C.J.F. terBraak, and O.F.R. vanTongeren. 1987. Data analysis in community and landscape ecology. Pudoc Wageningen Publishing. Netherlands.

Karr, J.R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications. 1:66-84.

Karr, J.R. and D.R. Dudley. 1981. Ecological perspectives on water quality goals. Environmental Management. 5:55-68.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and L.J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Special publication 5. Illinois Natural History Survey.

Kenkel, N.C. and L. Orloci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology. 67(4): 919-928.

Kerans, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications. 4:768-785.

Kruskal, J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika. 29:1-27.

Ludwig, J.A. and J.F. Reynolds. 1988. Statistical ecology: a primer on methods and computing. John Wiley and Sons, Inc. New York, New York.

McCune, B. and M.J. Mefford. 1997. PC-ORD. Multivariate analysis of ecological data, Version 3.0. MjM Software Design. Gleneden Beach, Oregon.

Mebane, C.A. 2001. Testing bioassessment metrics: macroinvertebrate, sculpin, and salmonid responses to stream habitat, sediment and metals. Environmental Monitoring and Assessment. 67:292-322.

Merritt, R.W. and K.W. Cummins. 1996. An introduction to the aquatic insects of North America. Third edition. Kendall/Hunt Publishing Co. Dubuque, Iowa.

Ohio Environmental Protection Agency. 1989. Addendum to biological criteria for the protection of aquatic life, volume II: user's manual for biological field assessment of Ohio surface waters. Ohio EPA, Division of Water Quality Planning and Assessment, Ecological Assessment Section. Columbus, OH.

Omernik, J.M. and A.L. Gallant. 1986. Ecoregions of the Pacific Northwest. EPA/600/3-86/033. U.S. EPA, Environmental Research Laboratory. Corvalis, Oregon.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers. 77(1): 118-125.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA 440-4-89-001. U.S. EPA, Office of Water Regulations and Standards. Washington, D.C.

Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris. 1995. Biological guidelines for freshwater sediment based on **BE**nthic **A**ssessment of **S**edimen**T** (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology. (1995) 20:198-219.

Robinson, C.T. and G.W. Minshall. 1998. Regional assessment of wadeable streams in Idaho, USA. Great Basin Naturalist. 58(1):54-65.

Royer, T.V., C.T. Robinson, and G.W. Minshall. 2001. Development of macroinvertebrate-based index for bioassessment of Idaho rivers: Environmental Management, v. 27, no. 4, p. 627-636.

Shannon, C.E. and W. Weaver. 1949. The mathematical theory of communication. The University of Illinois Press. Urbana, Illinois.

Simpson, E. H. 1949. Measurement of diversity. Nature 163(4148):688

Southerland, M.T. and J.B. Stribling. 1995. Status of biological criteria development and implementation. Pages 81-96 in W.S. Davis and T.P. Simon (editors). Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers. Boca Raton, Florida.

Stribling, J.B., B.K. Jessup, and J. Gerritsen. 2000. Developmental biological and physical habitat criteria for Wyoming streams and their use in the TMDL process. Prepared for U.S. EPA Region 8. Denver, Colorado.

Stribling, J.B., B.K. Jessup, J.S. White, D. Boward, and M. Hurd. 1998. Development of a benthic index of biotic integrity for Maryland streams. Maryland Department of Natural Resources, Report no. CBWP-EA-98-2 for the Maryland Biological Stream Survey. Annapolis, Maryland.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Aust. J. Ecol. 20:181-197.